Psychological Review
2007, Vol. 114, No. 1, 71–103

0033-295X/07/$12.00 DOI: 10.1037/0033-295X.114.1.71

Parallel and Serial Processes in Visual Search

Thomas L. Thornton and David L. Gilden
University of Texas at Austin

A long-standing issue in the study of how people acquire visual information centers around the scheduling and deployment of attentional resources: Is the process serial, or is it parallel? A substantial empirical effort has been dedicated to resolving this issue (e.g., J. M. Wolfe, 1998a, 1998b). However, the results remain largely inconclusive because the methodologies that have historically been used cannot make the necessary distinctions (J. Palmer, 1995; J. T. Townsend, 1972, 1974, 1990). In this article, the authors develop a rigorous procedure for deciding the scheduling problem in visual search by making improvements in both search methodology and data interpretation. The search method, originally used by A. H. C. van der Heijden (1975), generalizes the traditional single-target methodology by permitting multiple targets. Reaction times and error rates from 29 representative search studies were analyzed using Monte Carlo simulation. Parallel and serial models of attention were defined by coupling the appropriate sequential sampling algorithms to realistic constraints on decision making. The authors found that although most searches are conducted by a parallel limited-capacity process, there is a distinguishable search class that is serial.

Keywords: visual search, attention, computational models, simulation

How does the mind open up to acquire information? Arguably, this remains one of the central unresolved issues in modern psychology. With the rise of the information-processing paradigm in the social sciences, the issue of acquisition was reframed in terms of a one versus many distinction. Accordingly, the central problem of attention research became developing principled accounts of serial and parallel processing (Sternberg, 1966, 1975). For well over 2 decades, this dichotomy captivated much of the relevant psychophysical research and figured predominantly in numerous attempts to understand how attention operated during search (e.g., *feature-integration theory*: Treisman & Gelade, 1980; *guided search*: Wolfe, Cave, & Franzel, 1989). Although not always made explicit, the search problem was formulated in terms of single foveations and how attentional resources—and not the eyes per se—are directed. Unfortunately, it was soon realized (and slowly accepted) that the visual search method on which the early theories were built is inherently flawed—in short, the methods in use simply did not allow the serial–parallel distinction to be made (Townsend, 1972, 1974, 1990), provided that it existed in the first place (Eckstein, 1998; Geisler & Chou, 1995; Pashler, 1987; Palmer, 1995).

In the absence of a technology for deciding if a particular data set is produced by a serial or a parallel process, the distinction has slowly been abandoned as not being theoretically important or as being simply irrelevant to what an active visual system actually does. As it is obvious that the eyes do move around in search through natural scenes, the problem of what is learned in a glance is being displaced by investigations into eye movement strategies (Eckstein, Beutter, & Stone, 2001; Najemnik & Geisler, 2005; Rajashekar, Cormack, & Bovik, 2002; Rao, Zelinsky, Hayhoe, & Ballard, 2002; Tavassoli, van der Linde, Cormack, & Bovik, in press; Zelinsky, Rao, & Hayhoe, 1997). It is often the case in cognitive psychology that questions tied to a particular methodology are discarded as new perspectives and approaches arise, but here, the discarding of the serial–parallel distinction is premature and unwarranted. The distinction ought never to have been attached to the methods used in Treisman's (Treisman & Gelade, 1980) articulation of feature-integration theory. The basic questions concerning the deployment of attention in single fixations appear to be well posed and therefore should not be dismissed unless it can be shown that the construal of attention required by the serial–parallel distinction is either wrong or inconsistent.

Our approach to the serial–parallel decision problem involves extensive brute-force simulation of a multiple-target search (MTS) method (van der Heijden, 1975). This method was mentioned by Townsend (1990) as one possible resolution to the pitfalls of single-target search (STS). We have found that many search tasks once presumed to be serial are in fact better explained as capacity-limited, parallel processes. However, we have also found that contrary to popular belief, not all small set-size searches are parallel (Grossberg, Mingolla, & Ross, 1994; Humphreys & Muller, 1993; Humphreys, Quinlan, & Riddoch, 1989; Pashler, 1987)

Thomas L. Thornton and David L. Gilden, Department of Psychology, University of Texas at Austin.

Preparation of this article was supported by National Institute of Mental Health Grants R01-MH58606 and R01-MH065272. We thank James Townsend, Jeremy Wolfe, Kyle Cave, Llewyn Paine, and Laura Marusich for their helpful comments, suggestions, and emendations. Finally, we acknowledge S. Spear and A. Stah for their continued and unwavering support.

The simulation code is posted on David L. Gilden's faculty Web site: http://www.psy.utexas.edu/psy/FACULTY/Gilden/Gilden.html

Correspondence concerning this article should be addressed to either Thomas L. Thornton or David L. Gilden, Department of Psychology, University of Texas at Austin, 1 University Station A8000, Austin, TX 78712. E-mail: t.thornton@mail.utexas.edu or gilden@psy.utexas.edu

and that it is in fact possible to reliably resolve a class of search problems with serial element scheduling. We begin our discussion with a brief review of the search method that has historically been used to decide issues of processing style and, more lately, issues of processing efficiency (Wolfe, 1998a, 1998b).

Single-Target Visual Search

Despite its many problems, the standard visual search methodology continues to be the preferred tool for investigating the character of attentional limitation. It is simple to implement empirically, and the data patterns it generates have always had concrete theoretical meaning (Treisman, 1988; Treisman & Gelade, 1980) even if motivating theories turned out to be wrong or incomplete. In what is by far the most common version of the experiment, observers conduct speeded searches to determine whether a single target element is or is not present in a display of distractor elements. Typically, the average response time (RT) to make this decision is plotted as a function of the number of total elements (set size = 1 + number of distractors). Though errors are recorded, they are rare in most implementations and are usually examined only to insure that there is no simple trading of speed for accuracy. Throughout this article, we refer to this empirical approach collectively as the method of STS.

In STS, attentional limitation is inferred from the target-present latencies. If the target element is well camouflaged within its distractor field, search will be difficult and time consuming, and this means that RT will increase with set size. When this increase is substantial (usually a linear function of set size, with a slope of approximately 20–50 ms per additional item), the implication is that the search is difficult—attentionally demanding. A paradigmatic example of a difficult search task is hunting for a sideways T element target among mirror-flipped distractor Ts (Logan, 1994; Wolfe, 1998b).

There are also searches wherein the target is so obvious that it literally pops out. This happens in searches based on differences in color, motion, orientation, size, luminance, and other feature dimensions known to be processed early in visual cortex. When the target pops out, it makes no difference how many distractors are present, and in this limiting case, there is no cost associated with set size. In fact, flat RT functions were originally thought to define the class of parallel processes. It was this association that Townsend (1972, 1974) was clarifying with the introduction of limited-capacity parallel processes.

Shortcomings of Single-Target Search

There are problems with the single-target method, one theoretical, one practical. The theoretical problem is that set-size effects and RT functions with slopes greater than zero are not necessarily indicative of serial process. A parallel process might suffer costs of divided attention, and this would lead to serial-appearing RT functions (Townsend, 1990). In fact, there is no way to sort out seriality from capacity limitation by analyzing RT slopes.

The second major limitation of STS concerns experimental design and stimulus generation. Most experiments use a range of set sizes spanning anywhere from 1 to 4 elements at the lower end to 24 or more elements at the upper end. Although these ranges may provide a statistically reliable estimate of slope, they may also introduce visual artifacts such as masking and low acuity in the periphery (Carrasco & Frieder, 1997; Carrasco, McLean, Katz, & Frieder, 1998; Carrasco & Yeshurun, 1998; Geisler & Chou, 1995). These artifacts have been shown to undermine a clear interpretation of RT slopes.

Multiple-Target Visual Search

In recent years, several alternative methodologies that might be able to distinguish serial from capacity-limited parallel processes (Townsend, 1990; Townsend & Wenger, 2004) have been proposed. The MTS method, as implemented by van der Heijden (1975), is one such example. This method is an extension of STS that, under appropriate constraints, decouples capacity limitation from processing protocol. MTS augments the standard design by including trials with more than one target. The participant's task is the same as in STS—to indicate whether any targets are present. This extension generates key diagnostic conditions known as *pure-target trials*, in which every element in the display is a target. The pure-target trials are the critical feature in this design.

The potential benefits of this method (Townsend, 1990) lie in the following logic: If target-present RT decreases as pure-target number increases (i.e., there is a *redundancy gain*), then processing is parallel—if not, then processing is serial.[1] This follows because, for a serial process, the first item visited in a pure-target display will always be a target regardless of set size. As soon as this element is identified as a target, the search can terminate with a "target-present" response. Assuming that the average identification time of individual elements does not vary with set size (the standard assumption), serial models must predict flat pure-target RTs with set size. The practical implication is that whenever redundancy gains are observed in data, serial models can be ruled out in favor of limited-capacity parallel processing. Standard parallel models predict that RT should decrease with target number either owing to statistical considerations (*race gains*; Raab, 1962) or via spatial pooling of evidence across channels (see Miller, 1982).

The multiple-target method as originally discussed by van der Heijden (1975) involves only small set sizes, from one to three. In our work, we have added an additional element to potentially increase the resolution of redundancy gains. In the context of search experiments, four is not a large number. STS methods often

[1] Though a standard serial architecture does not predict redundancy gains, there are amendments to the basic model that can in principle achieve such effects. For example, a serial model with a favored position for processing (Egeth & Mordkoff, 1991; Mullin, Egeth, & Mordkoff, 1988; van der Heijden, La Heij, & Boer, 1983) can generate redundancy gains provided that there is a reliable benefit conferred by having a spatial bias (e.g., by reducing the time associated with orienting attention to that location, etc.) and that at least one element falls within this position. If these provisions are met, then on pure-target displays containing two or four targets, it is more likely that one of the targets will occupy the serial model's favored position (relative to Set Size 1). The more likely a target is to fall within the privileged position, the faster the average RT. In MTS, elements do not occupy fixed locations, and set size varies randomly, so that these kinds of favored-position effects are much less of a concern. Moreover, we have conducted simple analyses to check for favored-position artifacts (see Thornton & Gilden, 2001; van der Heijden et al., 1983), and we have found no evidence for such effects in any of our data.

use as many as 12 or 24 elements. We prefer a small cap on set size for several reasons. First, our studies are exploratory, and we wish to create a test bed with as few artifacts as possible. Larger set sizes introduce acuity artifacts as elements move into the periphery (constant density) or are packed more densely (constant area) (Geisler & Chou, 1995; Palmer, 1995; Palmer, Verghese, & Pavel, 2000), as well as promoting seriality through explicit eye movements. Also, as the element number increases, the overall pattern becomes more texturelike, and there is no question that textures have different grouping properties than individual elements (Wolfe, 1992; Wolfe, Chun, & Friedman-Hill, 1995). Nevertheless, inferences drawn from small element methods may not generalize to more complex search fields, an issue to which we return in our concluding comments.

Moving Beyond a Naïve Implementation of the Multiple-Target Search Method

Our initial interest in the MTS method arose from efforts to distinguish the different attentional demands placed by rotational and translational motions (Thornton & Gilden, 2001). In this work, we were particularly interested in the presence of redundancy gains in search for targets defined by direction sign. Three motion classes were investigated: translation (left, right), rotation (clockwise, counterclockwise), and looming–receding motion (translation along the line of sight). The data from these experiments are plotted in Figure 1. The average median RTs for correct trials only are plotted (upper graphs), with the associated error rates below them. In these plots, the target-absent data are connected by black, dashed lines, and the pure-target trials are denoted by the leftmost symbols on each of the one-, two-, and four-target tracks. We found that only rotation failed to show redundancy gains in the pure-target trials, implying that this motion uniquely requires a serial process for the acquisition of direction sign.

Although all of the target-present data are consistent with this interpretation, the target-absent data are problematic as it is not immediately clear what they signify. In all cases, the target-absent RT trends appear to mirror the pure-target trends (the only real difference between target-present and -absent RTs is in terms of a constant absolute offset in RT that is typical of yes–no paradigms). Obviously, the interpretation of the target-absent data cannot be the same as that of the target-present data even though they have the same shape. How is it possible to reject four distractors faster than one or two? Until redundancy gains in the target-absent conditions are understood, making inferences about processing style from the redundancy gains in the pure targets is simply not possible.[2]

Speed–Accuracy Trade-Off

As we gained further experience with MTS, it became increasingly clear that target-absent–pure-target mirroring is commonplace (Thornton, 2002). Typically, shallow or flat target-absent RT functions have been taken to indicate configural effects and the grouping of similar distractors (Humphreys, Quinlan, & Riddoch, 1989). In such cases, the target-present RTs are also typically flat or shallow with set size. This is a kind of mirroring, but we have seen mirroring even when there are large set-size costs (e.g., the rotation RTs in Figure 1). Additional investigations in our laboratory revealed that this kind of mirroring persists even when grouping and texture segmentation are attenuated (Thornton, 2002). Instead, the mirroring characterizing MTS data seems to be associated more with speed–accuracy trade-offs than with element grouping.

It appears that our observers were systematically scaling response criteria with set size in such a way as to reduce RT costs when targets were not present. Any adjustment of criteria that reduces RT will necessarily increase error rate, the miss rates in particular. Examination of the error-rate patterns in Figure 1 shows they are consistent with this idea. For all three motion experiments, there is a general trend for single-target misses to increase with set size, whereas false alarms hold constant or decrease with set size. This pattern of error is especially evident in the case of search for rotation direction, where misses for a single target among three distractors approach 20%. This particular pattern of rising miss rates and low false alarms has also been seen in some single-target experiments (e.g., Rensink & Enns, 1995) and may reflect the fact that observers were using a rational strategy (see Zenger & Fahle, 1997, for a treatment of error patterns in STS). A moment's thought suggests that observers might well truncate searches when targets are rare in the design (Wolfe, Horowitz, & Kenner, 2005)—especially when given instructions to respond quickly.

The standard approach for dealing with speed–accuracy trade-off is to either explicitly compute the trade-off function (McElree & Carrasco, 1999; Meyer, Irwin, Osman, & Kounios, 1988; Pachella, 1974; Wickelgren, 1977), construct an analytic function that combines RT and error (e.g., RT divided by accuracy; see Dennis & Evans, 1996; Townsend & Ashby, 1983), or constrain either RT or error so as to reduce its variation (Palmer, Ames, & Lindsey, 1993). None of these approaches is possible in MTS. Observers partition their error nonuniformly among the different treatment cells, and this is an unavoidable consequence of using a design that has variable numbers of targets and set sizes. The complexity of the error patterns is such that individual trade-off curves cannot be constructed, and we doubt whether it would be profitable to develop a separate trade-off rule for each cell. Instead, we have allowed explicit simulation of the search process to generate error partitions and have addressed speed–accuracy trade-off within the context of model fitting.

Construction of a Modeling Test Bed

Developing models of visual search processes requires a large test bed of data. These data should span the gamut of search

[2] The mirroring between conditions is puzzling because flat or decreasing target-absent RT functions are not predicted by any standard model of search (serial or parallel). Both standard serial and parallel models of processing predict that target-absent RTs should increase with set size. In the case of a serial process, this prediction is rather straightforward. Target-absent responses can only be made after all elements have been identified as nontargets, and thus, increases in set size necessarily lead to increases in target-absent RTs. In the case of a parallel process (capacity unlimited or not), the prediction is also that target-absent RTs should increase with set size purely because of statistical considerations (the slowest of *n* processes limits an exhaustive response).

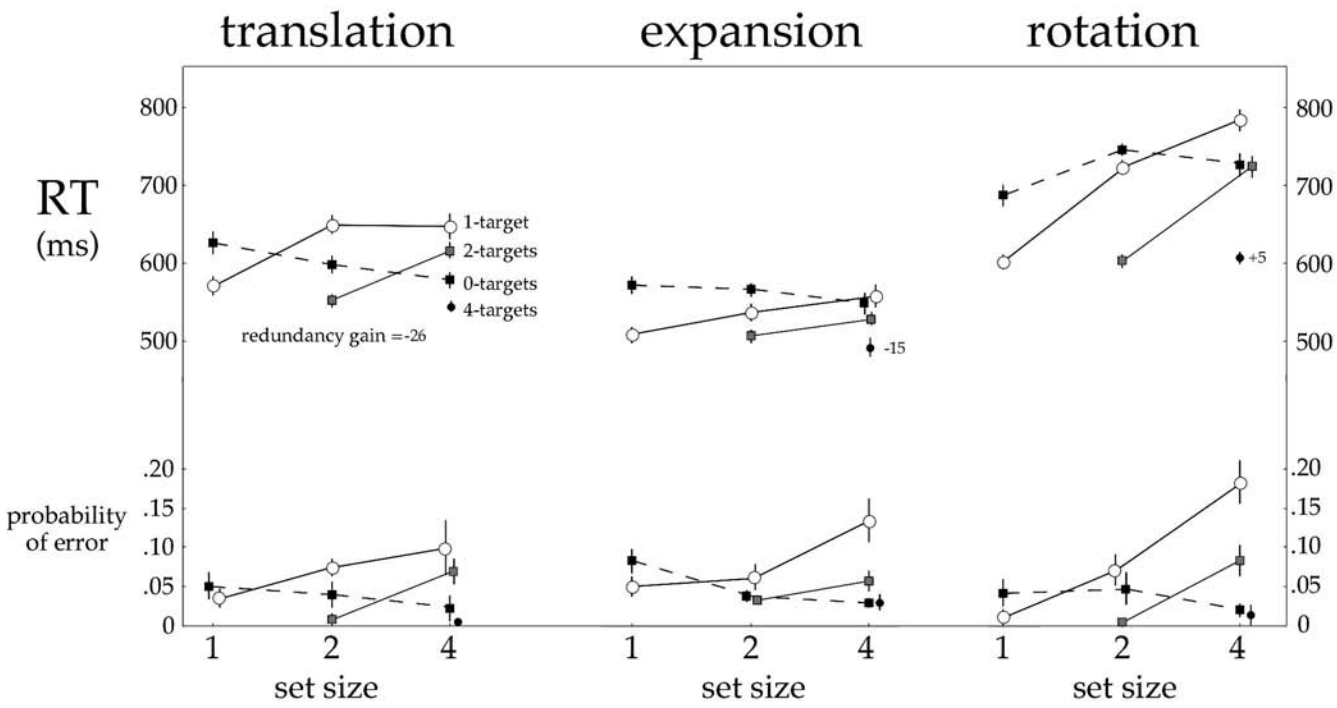

Figure 1. Multiple-target search (MTS) results for three motion-sign experiments adapted from Thornton and Gilden (2001). The upper graphs plot patterns of average response time (RT) for each of the nine conditions in MTS; the lower graphs plot the associated average error rates (error bars denote standard errors). Nine observers participated in the translation and expansion experiments; 18 observers participated in the rotation experiment.

difficulty so that the models can be calibrated across the full range of attentional demand. In this section, we describe the 29-study test bed we created, the core methodology, and all pertinent details for replications of these studies.

In addition to the motion tasks that motivated our original studies, the remainder of tasks in the test bed were chosen first to represent the continuum of search difficulty and second to represent the key stimulus sets that have historically shaped theory development in this field. The ensemble includes tasks that must support parallel processes, such as search for color, as well as tasks that are known to be challenging, such as looking for targets that differ from distractors by a mirror reflection.

Method

For all MTS measurements, displays contained either one, two, or four elements. Individual elements were configured about a central fixation point along a virtual circle whose radius varied from 1.5° to 2.5° of visual angle. Elements were drawn at canonical locations along the virtual circle (45°, −45°, 135°, −135°), and for most tasks, the entire display was randomly rotated about fixation to remove configural effects by choosing a uniform deviate from the interval ±25°. For some of the tasks, additional radial jitter (~0.5°) was added individually to each element to remove effects due to element colinearity. A schematic of the general protocol for display generation is shown in Figure 2A.

All target–distractor differences were of high contrast, and we verified on numerous occasions that all elements were both highly discriminable and categorizable. The majority of individual elements across all of the studies subtended about 1° or 2° at a viewing distance of 57.3 cm. In all, there were nine basic types of stimulus displays; displays containing three targets were excluded from this design. Displays consisted of all distractors (target-absent trials), all targets (pure target-present trials), or some combination of a variable number of targets and distractors (mixed target-present trials). Figure 2B shows the relative probabilities of encountering each type of stimulus display. This particular design matrix was necessary to insure that the probability of encountering target-present and target-absent displays was balanced across set size.

Nine different observers participated in each motion-sign experiment, except for Task 26 (rotation textures), where data were pooled over an additional replication to yield 18 observers. Eight observers participated in all other experiments, except for Task 2 (orientation), where data were again pooled over an additional replication to yield 16 observers. For all tasks, stimulus displays were preceded by a brief fixation interval (~500 ms) and were present until response.

Stimuli

In total, the test bed consisted of 29 search tasks, each of which may be broadly categorized into one of six groupings: (a) *featural,* (b) *emergent cues,* (c) *rotation-induced distractor heterogeneity,* (d) *conjunction,* (e) *configuration,* and (f) *rotations*. These groups summarize our impressions of task similarity and do not at this

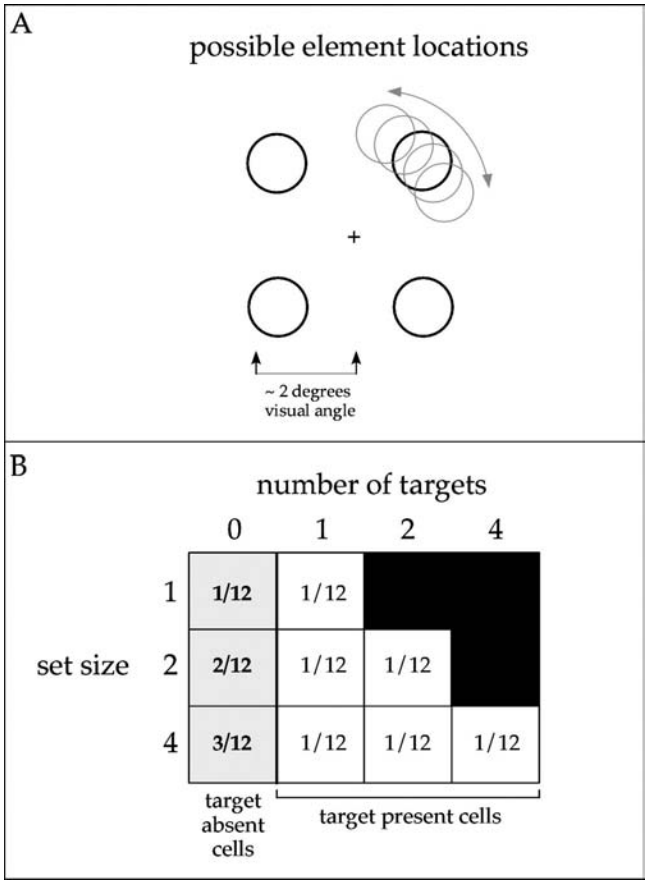

Figure 2. Details of the multiple-target search (MTS) method. A: Schematic of the stimulus arrangement protocol used for all tasks. B: Relative probabilities of the nine trial types used in MTS. Each trial type is defined by one of three set sizes (1, 2, or 4) and one of four target numbers (0, 1, 2, or 4); the number of distractors present in any display equals the set size minus the number of targets. Probabilities for the target-absent conditions are shown in the leftmost column (gray squares) and increase with set size so as to balance the probability of target-present and target-absent trials within set size.

point confer any theoretical commitment. For reference, Figures 3, 4, 5, 6, 7, and 8 show examples of the stimulus displays that were used in each of the 29 tasks. In every panel, there is an example of a single-target, three-distractor display (the element assigned as target appears in the upper left quadrant).

Featural Search

The following eight tasks consisted of searches based on targets and distractors that differ along a single feature dimension. These are the classic pop-out searches and are well known to generate spatially parallel search and vivid texture segmentation. The looming–receding motion task is an oddball in this regard because this aspect of motion direction is encoded first in the dorsal division of the medial superior temporal area (Graziano, Anderson, & Snowden, 1994; Tanaka, Fukada, & Saito, 1989; Tanaka & Saito, 1989), whereas the other features generate responses in visual cortical area 1. Also, unlike the other feature contrasts in this group, looming and receding motion patches do not generate strong texture segmentation (Gilden & Kaiser, 1992). We include the task in our featural group only because recent visual search studies suggest that it is acquired efficiently and in parallel (Takeuchi, 1997; Thornton & Gilden, 2001). Key members of this class serve to benchmark and establish the external validity of our models. The models must find, for example, that suprathreshold color differences are processed in parallel—because they are in fact processed in parallel. Example stimulus displays from each task are shown in Figure 3.

Task 1: Color. Elements were circular, colored disks (~1.33° visual angle); targets were reddish gray, and distractors were bluish gray; elements were made highly similar by reducing saturation—this was done primarily to protract overall RTs (in an earlier pilot study using highly discriminable color differences, the pure-target RTs were so fast as to preclude any observable redundancy gains).

Task 2: Orientation. Elements were windowed sinusoids (sine-phase Gabors, 2.5 cycles per degree, with a 1.5° Gaussian envelope). Targets were oriented 45° left of vertical; distractors were vertical.

Task 3: Size (big target). Targets were large, low-frequency Gabors (2.5 cycles per degree, with a 1.5° envelope); distractors were small, higher frequency Gabors (4 cycles per degree, with a 0.75° envelope).

Task 4: Size (small target). This task used the same stimuli as Task 3 with target and distractor roles reversed.

Task 5: Translation. Elements were continuously translating naturalistic textures moving behind circular apertures (~3° visual angle); targets moved rightward, and distractors moved leftward (this experiment was previously reported in Thornton & Gilden, 2001).

Task 6: Expansion. Elements were identical to Task 5 but consisted of expanding–contracting texture with realistic two-dimensional acceleration; targets were expanding, and distractors were contracting.

Task 7: Contraction. This task used the same stimuli as in Task 6; target and distractor roles were reversed (data from Tasks 6 and 7 were reported previously in Thornton & Gilden, 2001).

Task 8: A among Bs. This task was included to investigate efficient letter search in the context of MTS. This particular search should come out parallel given that the target and distractor letters differ from each other along a number of basic stimulus dimensions (e.g., local orientation, curvature, closure, terminations, etc.). Both letters were uppercase (~1.3° visual angle) and were white on a black background; targets were As, and distractors were Bs. The letter elements were ramped up from zero to full contrast over the course of a second (this was done to protract the very fast RTs observed in pilot studies).

Emergent Cues

This class is particularly interesting in that the features that distinguish targets from distractors are standard examples of perceptual organization. Shape from shading (Ramachandran, 1988), the use of perspective to generate depth (Enns & Rensink, 1990), and the famed parenthesis objects (Pomerantz & Pristach, 1989)

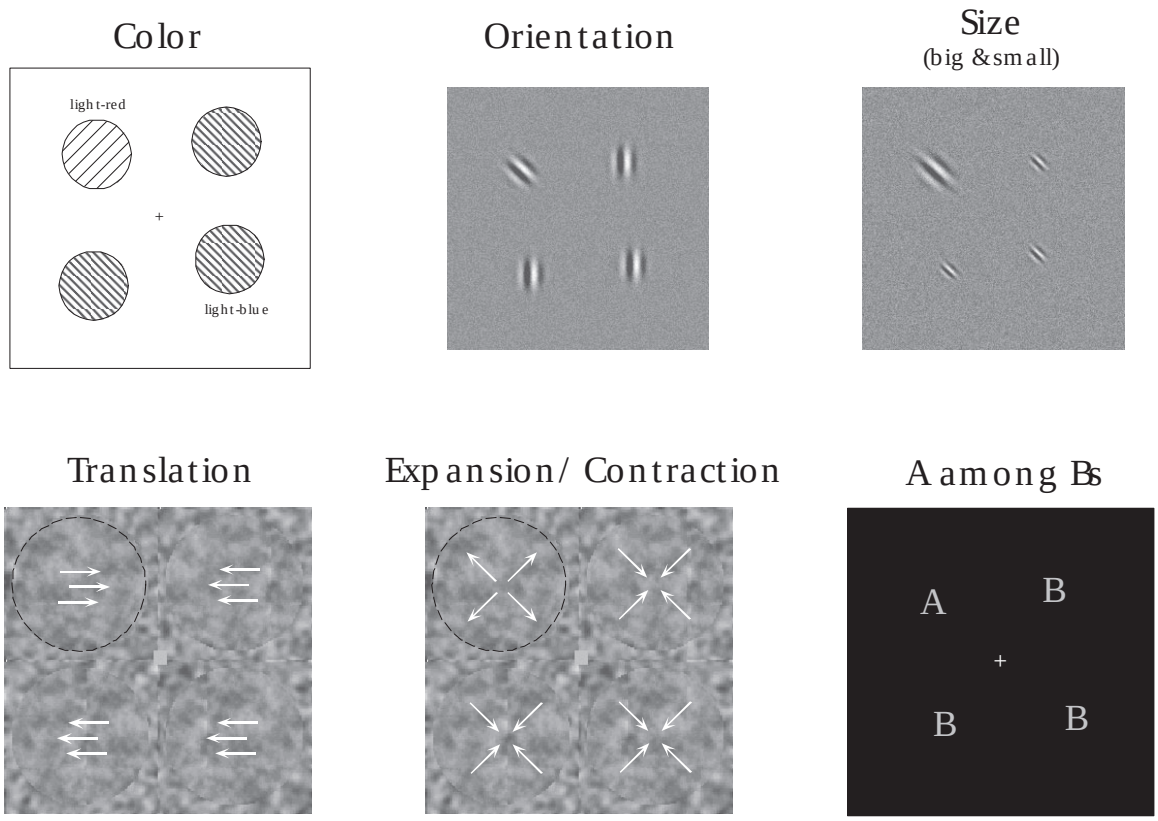

Figure 3. Example stimulus displays from the eight tasks that make up the featural group (color, orientation, size-big, size-small, translation, expansion, contraction, and A among Bs). Each display shows a Set Size 4, single-target condition (the targets are consistently the upper left elements in each display).

are in this group. The common wisdom based on STS is that depth and emergent shape may lead to highly efficient searches, and this is interesting because these contrasts are not based on simple features. The additional degrees of freedom in MTS clarify the extent to which this is true. Example stimulus displays from this class are shown in Figure 4.

Task 9: Shading-UD. This task examined search for emergent shape based on shading and an assumed lighting source (Ramachandran, 1988; Sun & Perona, 1996). Elements were circular disks (~1.5° visual angle) with a shading gradient running from top to bottom; targets were shaded from white to black and were perceived as surface bumps; distractors had opposite polarity and were perceived as surface dimples. Small reverse-shaded inducer elements were included so as to increase the percept of shape in Set Size 1 displays.

Task 10: Boxes. This stimulus set was based on work by Enns and Rensink (1990) showing that an implied organization in depth (using only two-dimensional elements) can generate efficient search. Elements shared a set of three colored faces in differing arrangements; targets were boxes that were white on top and pointed up and to the left; distractors were white on the bottom and pointed down and to the right.

Task 11: Parentheses. This task used search stimuli similar to those originally investigated by Pomerantz and Pristach (1989). Elements were distinguished by configurations of a set of two curved lines and were constructed so as to minimize any difference in overall spatial extent; targets had the curves facing each other so as to create an ovallike group, whereas distractors had the curves facing out to create an hourglass group.

Rotation-Induced Heterogeneity

The following four search tasks were based on shape discrimination and were purposely made difficult by distractor heterogeneity. This was achieved by randomly rotating all of the elements. The first two tasks consisted of search for a T among Ls, which is of particular interest given the general consensus in the literature that these stimuli demand a serial analysis (Bergen & Julesz, 1983; Egeth & Dagenbach, 1991; Treisman & Gelade, 1980; Wolfe et al., 1989). The final two tasks examined search for triangles presented among a set of diamonds (Task 14) or among a set of polygons (with varying numbers of sides; Task 15). Distractor heterogeneity is known to increase attentional demand (Duncan & Humphreys, 1989), but it is not known whether random rotations induce serial process. Stimulus examples appear in Figure 5.

Task 12: TL-white. Elements were letters (~1.8° visual angle) presented on a black background and were constructed from a shared set of three line-elements (this was done to equate compo-

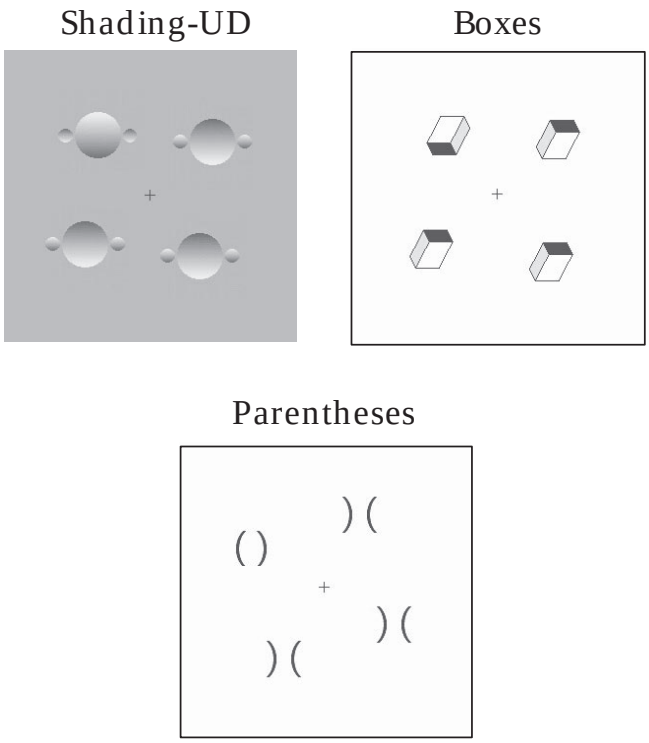

Figure 4. Example stimulus displays from the three tasks in the emergent features group (shading-UD, boxes, and parentheses).

nent orientation, length, and overall luminance across target and distractor). Targets were Ts, and distractors were Ls; all elements were randomly given an orientation of 0°, 90°, 180°, or 270°.

Task 13: TL-black. This task provided a replication of Task 12 using smaller, black letters (~0.8° visual angle) on a white background.

Task 14: Triangles. Elements were simple black polygons that were randomly rotated to remove cues based on local orientation; targets were triangles, and distractors were diamonds. All shapes were constructed so as to minimize any perceived distinctions in spatial extent.

Task 15: Polygons. Elements were similar to those used in Task 14. Targets were randomly rotated triangles, and distractors were randomly sampled from a heterogeneous set of three polygons (a pentagon, diamond, or hexagon matched in phenomenal size to the target).

Conjunction Search

We have conducted two variants of this classic search task in which target and distractor elements are defined via a conjunction of feature values. The presumed serial nature of these kinds of tasks has figured prominently in tests of feature-integration theory (Treisman & Gelade, 1980) and guided search (Wolfe et al., 1989) and continues to play an important role in more recent work on signal-detection models of search (e.g., Eckstein, 1998; Eckstein, Thomas, Palmer, & Shimozaki, 2000). By definition, conjunction search tasks have heterogeneous distractors. The theoretical issues surrounding conjunction search are essentially whether this form of heterogeneity is sufficient to require serial process. Examples of the stimulus displays used in these tasks are shown in Figure 6.

Task 16: Color-orientation. This particular conjunction task utilized a Color × Orientation stimulus set known to consistently yield moderate to large set-size effects in STS (Nakayama, Wang, & Kristjansson, 2000; Wolfe, 1998a). Targets were white verticals; distractors shared one feature with the target and consisted of black verticals and white horizontals. Displays were constructed by randomly sampling distractors, with the constraint that homogeneous distractor draws were not allowed (without this constraint, targets can be distinguished solely by color or orientation).

Task 17: Bigram-conjunction. This task examined conjunction search in the context of simple letter-strings. The stimuli were drawn from a previous set of studies that used STS data to argue for serial processing of words (Duncan, 1989). The search elements consisted of strings of four black letters (~1° visual angle) presented on a white background. The target was the string STAB, consisting of the concatenation of the bigrams ST and AB. The distractors consisted of random draws from a set of two strings that shared one bigram with the targets (STUX, ICAB). In Set Size 4 displays, distractor heterogeneity was maintained so that the target could not be distinguished solely by a unique letter.

Configuration

The following tasks examined the most difficult searches—when targets and distractors can only be distinguished in terms of

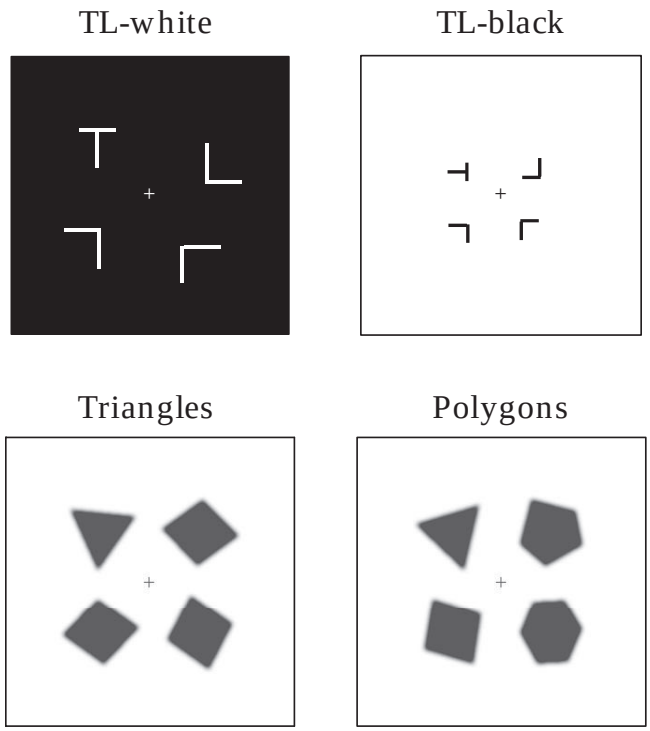

Figure 5. Example stimulus displays from the four tasks in the rotation-induced heterogeneity group (TL-white, TL-black, triangles, and polygons).

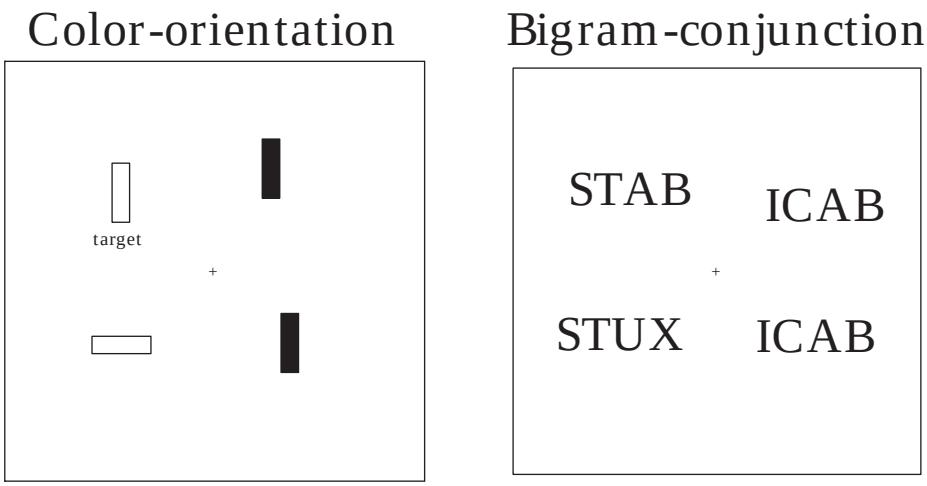

Figure 6. Example stimulus displays from the two tasks in the conjunction group (color-orientation and bigram-conjunction). In both cases, distractors were randomly sampled from a two-element set with the additional constraint that, on any given trial, the distractors could not all be identical.

the relative position or configuration of shared components. There is a large body of work attesting to the difficulty subtle changes in configuration impose on attention. This has been documented in STS (Enns & Rensink, 1990; Logan, 1994; Moore, Egeth, Berglan, & Luck, 1996; Saarinen, 1996; Wolfe, 1998a; Wolfe & Bennett, 1996), threshold search based on accuracy (Palmer, 1994; Poder, 1999), and studies of texture segmentation (Beck, 1966; Geisler, Stern, Thornton, Kuyel, & Ghosh, 1998; Malik & Perona, 1990; Rentschler, Hubner, & Caelli, 1988; Sagi, 1995). From our point of view, these tasks are key members of the test bed because the models must reflect the reality that these kinds of search are attention demanding. Either the models must assign large costs for dividing attention in parallel or they must fit the data with a serial process. Examples of the stimulus displays used in these measurements are shown in Figure 7.

Task 18: Missing-side. Elements were highly discriminable black C-like figures (~1.2° visual angle) constructed from a shared set of three lines; targets had a missing side on the right, whereas distractors had a missing side on the left.

Task 19: Y-UD. Elements were Y junctions (~1.3° visual angle) oriented upward for targets and downward for distractors.

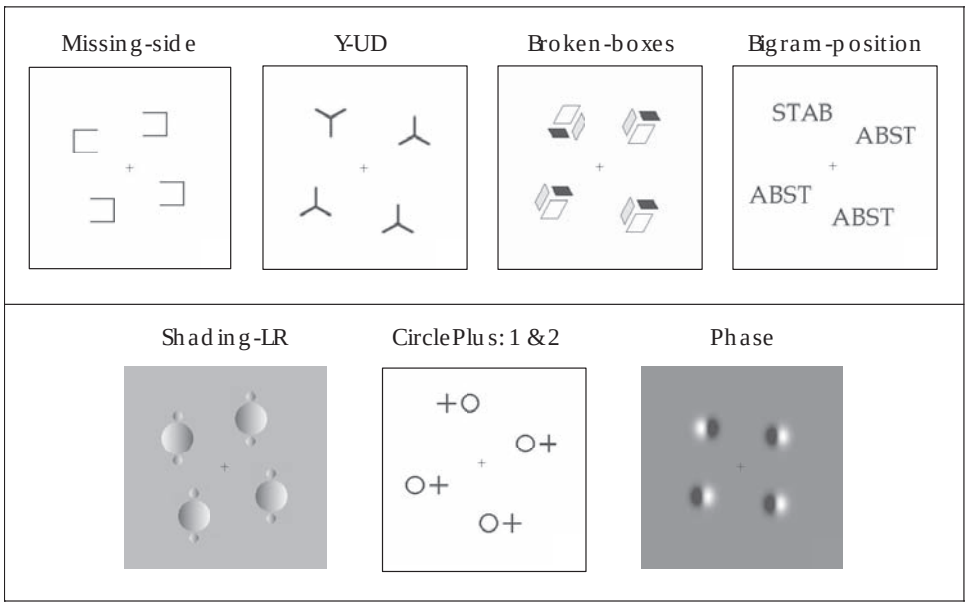

Figure 7. Example stimulus displays from the eight tasks making up the configural group (missing-side, Y-UD, broken-boxes, bigram-position, shading-LR, circlePlus-1, circlePlus-2, and phase).

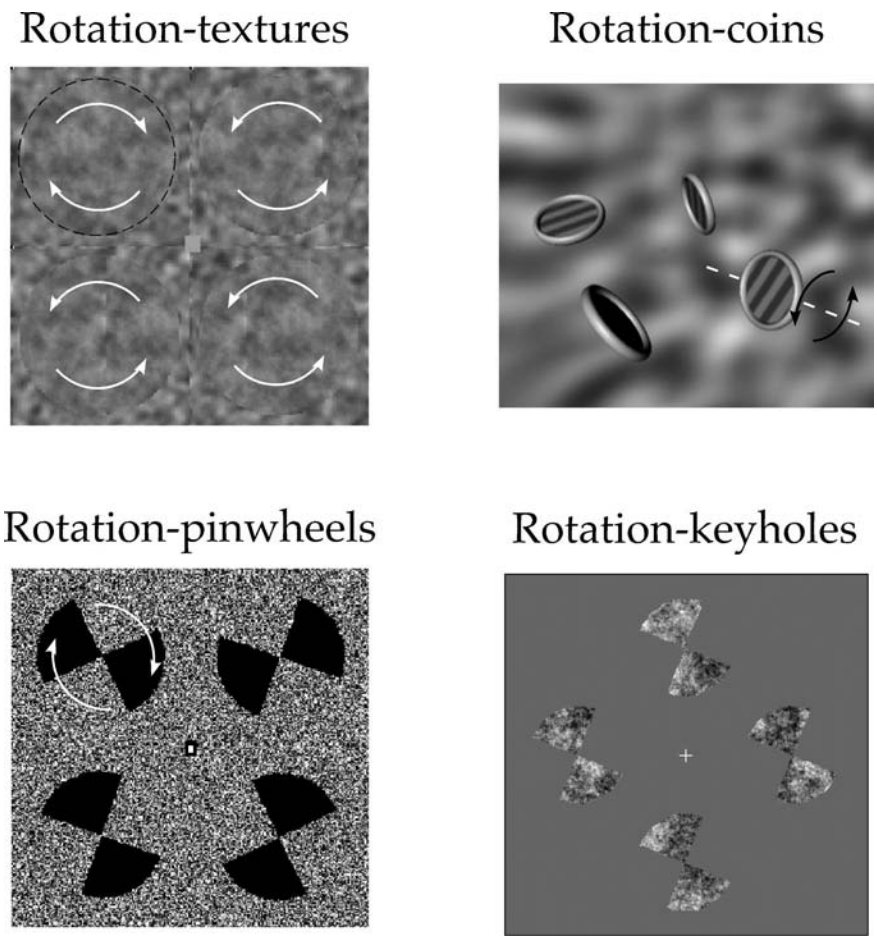

Figure 8. Example stimulus displays from the four tasks in the rotations group (rotation-textures, rotation-coins, rotation-pinwheels, and rotation-keyholes).

Task 20: Broken-boxes. These elements were based on the same stimulus sets used in Task 9 (boxes; see also Enns & Rensink, 1990) except that the colored, trapezoidal faces were separated so as to attenuate the three-dimensional cubelike interpretation. With this change, targets can only be distinguished from distractors by a particular configuration of trapezoids.

Task 21: Shading-LR. This task was based on the same stimulus sets used in Task 8 (shading-UD) except that the target and distractor elements were rotated 90° counterclockwise. This manipulation removes distinctions in surface curvature and reduces the task to a strict judgment of shading polarity.

Task 22: CirclePlus-1. Elements consisted of arrangements of a circle and a plus sign (~1.2° × 2.2° visual angle) presented on a white background; targets had the plus to the right of the circle, whereas distractors had the reverse arrangement. Though these stimuli are highly discriminable psychophysically, in the context of multielement displays they are known to generate inefficient patterns of search characterized by large set-size effects (Logan, 1994; Moore, Elsinger, & Lleras, 2001).

Task 23: CirclePlus-2. This task served to replicate Task 22 with a different set of observers.

Task 24: Phase. Elements were one-cycle square waves blurred by a circular Gaussian envelope (~1.3° visual angle)—targets were white on the left, and distractors were white on the right. These stimuli were similar to those used in Task 21 but were of higher contrast, had a sharper gradient, and did not include inducer elements.

Task 25: Bigram-position. This stimulus set examined search for a specific configuration of shared bigrams (Duncan, 1989). Elements were identical to those used in Task 17 with the following change—the distractors were no longer heterogeneous and were defined by a reversed arrangement (ABST) of the bigrams used in the target string (STAB).

Rotations

The following four tasks examined search for a specific direction of rotation. In contrast to translational motion, the sensing of rotation direction generates seriallike patterns in search (Thornton & Gilden, 2001) and does not permit effortless texture segmentation (Julesz & Hesse, 1970). One of the benefits of reassessing this group is that we were able to replicate our earlier findings and to

put them on a more solid basis with the modeling approach that was developed subsequent to the original studies. Examples of the stimulus displays used in these tasks are shown in Figure 8.

Task 26: Rotation-textures. Elements were continuously rotating naturalistic textures moving behind circular apertures (~3° visual angle)—targets rotated clockwise, and distractors rotated counterclockwise (this experiment was previously reported in Thornton & Gilden, 2001).

Task 27: Rotation-coins. This task extended the investigation of rotation search to the case of motion about an axis oriented in depth. Elements were realistic coinlike objects (rendered in Cinema 4D; Maxon, Newbury Park, CA). Targets and distractors rotated in different directions about an obliquely oriented axis. Unlike the stimuli used in our other motion tasks, individual elements differed slightly in size within a display (owing to perspective projection) and were given realistic lighting cues and specularities to strengthen the three-dimensional percept.

Task 28: Rotation-pinwheels. Measurements on this task were previously reported in Thornton and Gilden (2001). The task examined rotation search when elements dynamically accreted and deleted background texture. This manipulation gave rotation a feature common to all translational displacements and tested whether this property was sufficient to generate efficient search for direction. Elements were black pinwheels (~3° visual angle) that rotated against a static Gaussian noise background (targets rotated clockwise, distractors counterclockwise).

Task 29: Rotation-keyholes. This task also examined rotation search in the context of accretion and deletion. In this variant, naturalistic textures rotated behind keyhole apertures (~3° visual angle), and object texture was accreted or deleted in a manner similar to the aperture-bounded translation and expansion–contraction stimuli used in Tasks 5–7.

Preliminary Data Reduction

For all 29 tasks, observers completed 288 trials of practice before providing either 576 trials (motion-sign tasks) or 864 trials (the remaining 22 tasks). In the preparation of the RT data, we excluded all trials on which errors occurred and trials with RTs greater than 1,500 ms or less than 150 ms. In no case was more than 1% of the data excluded (this did not vary across the test bed). Prior to averaging, we converted RT data from individual observers to *z* scores using each observer's global mean and standard deviation.

Transformation to *z* scores before averaging across observers has a number of desirable properties. First, a *z*-score transformation uses each observer's intrinsic variability to bring all the RT effect sizes onto a common scale prior to averaging—a 100-ms set-size effect is phenomenally much larger for an observer whose overall standard deviation is 50 ms than it is for an observer whose standard deviation is 200 ms. Second, it removes individual-differences variability arising from overall speed—some people are faster than others. The variability that is left and that is relevant to model selection arises from how participants differentially effect speed–accuracy trade-offs across the nine cells of the design. In this way, it is the patterns of RT data that are highlighted in the analysis, not where the patterns happen to fall in terms of an absolute number of milliseconds. The models indeed make no prediction about how many milliseconds any aspect of search will consume. Removing between-participant overall speed differences effectively shrinks the error bars in RT patterns, providing a more demanding test for models in goodness-of-fit comparisons.

After normalization to *z* scores, we computed within-observer RT medians for each of the nine search conditions. Median *z* scores were used to compute cell means and standard errors. Models were fit to the averaged data using a resampling procedure so that error bars on the model parameters could be estimated. It should be noted that although the RT distributions collected from individual participants are nonnormal (and so too are the derived *z*-score distributions), the data that we modeled are averages over observers and so are approximately normal. We used the normality of these distributions to obtain confidence limits on parameter estimation (see Appendix A, section entitled Resampling Data).

For clarity of presentation and intertask comparison (targets in different experiments are more or less easy to find), we converted the final averaged RT *z* scores and standard errors back to units of seconds using each task's global RT mean and standard deviation (averaged over observers). This last step was taken simply to present our results in dimensional units and in no way influences the fitting and development of models.

Insofar as RT patterns in visual search cannot be understood outside of the context of speed–accuracy trade-offs, we have found it necessary to also compute averaged error-rate patterns for each task. Because the errors tended to have a common range across all observers and tasks (a floor at 0 and less than 10% errors on average), no *z*-score transformation was used here prior to averaging over participants. These two sets of patterns, one in RT, one in error, were jointly used to select models.

RT and Error Patterns in Multiple-Target Search

We found three basic patterns of data, illustrated in Figure 9, that were reiterated throughout the entire test bed, and they fairly describe what might be expected in MTS. Eventually, we identify these patterns with specific models of visual search. Here, we introduce the patterns in terms of their most salient and identifying features. The individual plots for each experiment are displayed in Appendix B and are grouped according to classes outlined in Figure 9.

Class A

1. Mild set-size effects in single-target RT and miss rates (average increases of approximately 43 ms and 6% error at Set Size 4);
2. Target-absent and pure-target RTs that mirror each other, decreasing similarly with set size (–27 ms and –17 ms, respectively); and
3. False alarms that decrease over set size by about 3%.

Class B

1. Set-size effects in single-target RTs roughly twice those seen in Class A (~89-ms increase at Set Size 4),

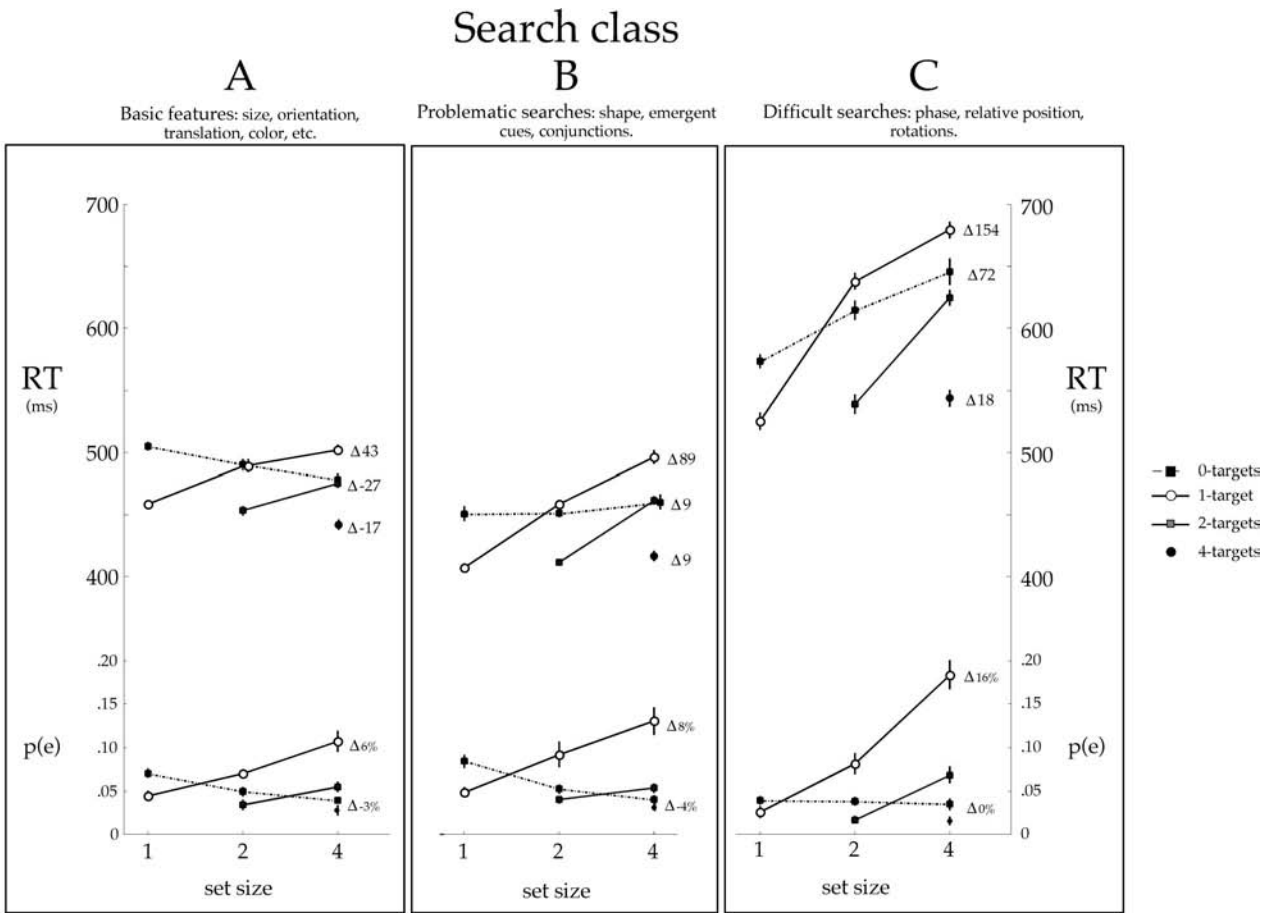

Figure 9. Three classes of data as revealed in the use of multiple-target search. The graphs in Panel A show the average response time (RT) and error data from a subset of 10 tasks based on simple featural distinctions. The graphs in Panel B show the average data from a subset of 11 tasks that have proven difficult to interpret in previous treatments; this class includes searches for shape, conjunctions, and emergent properties. The graphs in Panel C show the average data from a subset of 8 highly demanding forms of search based on relative position, phase, and rotation direction. The inset numbers in the three upper graphs denote the RT differentials over set size. From top to bottom, the values correspond to (Set Size 4, single-target RT) minus (Set Size 1, single-target RT), (four-distractor RT) minus (Set Size 1, single-distractor RT), and (four-target RT) minus (Set Size 1, single-target RT), respectively. The inset numbers in the lower three graphs denote error-rate differentials over set size. From top to bottom, the values correspond to (Set Size 4, single-target miss rate) minus (Set Size 1, single-target miss rate), and (Set Size 4, false-alarm rate) minus (Set Size 1, single-distractor false-alarm rate). Error bars denote standard error of the mean.

2. Slightly larger increments (relative to Class A) in single-target misses (approximately 8% at Set Size 4),
3. Target-absent and pure-target RTs that mirror each other with no appreciable redundancy gains in pure-target RT and similarly flat target-absent RT patterns, and
4. False alarms that decrease over set size by about 4%.

Class C

1. Low single-element errors with little observable RT asymmetry (both misses and false alarms around 2% to 3%);
2. Large set-size effects, whereby single-target RTs and misses rise dramatically (154 ms and 16%, respectively, in going from zero to three distractors);
3. Target-absent RTs that rise with set size at roughly half the rate of single-target RTs (an approximately 72-ms increase);
4. Pure-target RTs that are flat or rise mildly with set size (on average, ~18 ms); and
5. False-alarm rates that are generally low and roughly constant over set size.

Class A tasks show the hallmark signatures of spatial parallelism in MTS in that set size has a relatively small influence on RT

and miss rate and there are strong redundancy gains in the pure-target conditions (i.e., RT decreases with target number). That these tasks generate clear evidence of parallel processing validates our methods because we know from physiology and psychophysics that distinctions in size, color, orientation, and so on are realized via massively parallel representations in early cortex (Kandel, Schwartz, & Jessell, 2005, provide an overview of these issues). Moreover, these tasks allowed us to calibrate the search methodology by defining an intrinsic ruler with which to measure other, more demanding tasks.

Class C tasks are known to be highly demanding and were included in the ensemble as an additional calibration point. The large single-target slope implies a high degree of attentional demand, and the pure-target losses with set size are at face value inconsistent with parallelism—even if capacity limited. Four of the eight tasks in the class are based on search for a unique direction of rotation. This particular stimulus distinction has been shown to yield seriallike data in previous work (Gilden & Kaiser, 1992; Julesz & Hesse, 1970; Thornton & Gilden, 2001). The remaining four tasks in Class C require judgments of relative phase or configuration, a similarly demanding distinction that generates highly inefficient search patterns in single-target treatments (Logan, 1994; Moore et al., 2001; Poder, 1999; Wolfe, 1998a, 1998b).

Class B contains data from search tasks that have historically been the most difficult to characterize. One of the principal aims of this work was to give this class proper definition. Class B data do not show redundancy gains in the pure-target trials, and this makes them look serial. However, the error patterns look more like those of Figure 9A, and Figure 9A must be exemplary of a parallel process given the featural tasks it includes. Class B includes conjunction search as well as search for specific emergent cues (boxes drawn in perspective, curvature based on implied lighting, gestalt closure). It has never been clear how to categorize conjunction search (Duncan, 1989; Eckstein, 1998; Treisman & Gelade, 1980; Wang, Kristjansson, & Nakayama, 2005; Wolfe et al., 1989; see also Wolfe, 1998a, and the many references therein). Furthermore, although it is established that emergent cues may make search more efficient (Enns & Rensink, 1990; Pomerantz & Pristach, 1989), set-size costs have never been estimated within a consistent paradigm and within a modeling approach.

In reality, Figure 9B is not unique in posing problems of interpretation. None of the data displayed in Figure 9 are entirely straightforward. The speed–accuracy trade-offs that distorted the target-absent RTs in our original studies (see Figure 1) are generic to the entire test bed. To our knowledge, there are no models that predict that people will be faster on target-absent trials with increasing set size.[3] Moreover, it is fundamental to the literature on error analysis that error rates increase whenever the opportunities for making errors increase. Yet we have never found increasing false-alarm rates with set size. Inferences about processing style must take into account trends in both RT and error, and these trends are simply too complicated to understand through informal visual inspection. To this end, we have modeled the data using Monte Carlo techniques that allow us to flexibly simulate the speed–accuracy trade-offs that create these counterintuitive trends. To the extent that we can adequately capture the patterns used by observers in MTS, we will be able to decide the parallel–serial problem for the three major classes illustrated in Figure 9.

Models

In the following section, we discuss the various issues that arise in the formulation of models of visual search. This discussion involves several parts: how attention and perception are represented, how decisions are represented, and how the model generates the requisite data—reaction times and error rates. The desired outcome is a formal decision procedure that stipulates the kind of process that most likely is responsible for the observed patterns of data.

Sequential Sampling and Diffusion-Based Evidence Accumulation

The core component of our modeling approach derives from a tradition of sequential sampling models that construe information pickup as a process of noisy evidence accumulation over time (Laming, 1968; Link, 1975; Ratcliff & Smith, 2004; Stone, 1960; Townsend & Ashby, 1983). Sequential sampling algorithms mimic the physics of diffusion to a boundary and represent an extension of signal-detection theory over time.[4] Their application in psychology has proven enormously successful in the modeling of two-choice decision (Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff

[3] In some of our earliest attempts to explain the target-absent RT patterns observed in MTS, we explored the utility of a model based on the notion of discriminating display homogeneity. The idea here was that target-absent and pure-target RTs had symmetric trends because observers were making fast initial judgments of homogeneity (i.e., whether all elements were of one type), followed by a more refined assessment of identity (target or distractor?). Though this kind of two-stage model does succeed in qualitatively explaining the mirroring seen in some data sets—in particular, those cases in which target-absent RTs are invariant or decrease with set size—we abandoned it as a complete model of MTS data for two reasons. First, without a number of additional assumptions regarding how perceived display heterogeneity scales with set size and target number, the homogeneity model makes incorrect predictions for mixed-element displays (e.g., the standard displays of STS). As soon as a mixed-element display is perceived to be inhomogeneous (via the first-stage analysis), then the model necessarily generates a "target-present" response with no need for a further analysis of individual element identity. By this logic, mixed-element displays should be analyzed faster overall and should have trends that track those of homogeneous displays (i.e., the pure-target and target-absent RT trends)—both predictions are at odds with actual MTS data. Second, although the central construct of homogeneity is easy to articulate, it is difficult to instantiate formally in a process model, especially if we assume it proceeds prior to element identification.

[4] In signal detection (Green & Swets, 1988), decision is modeled as the comparison of a single sample of noisy evidence with a fixed decision bound (if the sample exceeds the bound, say signal, else noise). In the most general version of a sequential sampling model, the sampling–comparative process of signal-detection theory is reiterated so that at each point in time, a new sample of evidence is added to one or more registers. The current value of total evidence in the register(s) is then compared with one or more decision bounds. Samples continue to accrue until enough evidence has accumulated to reach one of the competing response boundaries. A number of other sequential sampling variants possess similar properties and so represent equally plausible alternatives for the modeling of multiple-target visual search (e.g., the Poisson race model: Pike, 1973; Townsend & Ashby, 1983; Van Zandt, Colonius, & Proctor, 2000).

& Smith, 2004, and the references therein). For our purposes, we used an approximation of continuous diffusion based on the random walk model of decision (Laming, 1968; Link, 1975).

There are several virtues that recommend diffusion models. Their properties are well understood mathematically (Link, 1975; Townsend & Ashby, 1983), and we are able to check our simulations against key analytic results. This provides crucial tests for debugging code. Diffusion models are also perfectly suited for simulating search because the termination of a random walk simultaneously yields a stopping time (reaction time) and a boundary absorption (correct or incorrect decision). The geometry of noisy diffusion to a boundary naturally leads to positively skewed distributions of performance variables, a feature common to empirical RT distributions (Luce, 1986). Most importantly, diffusion models allowed us to graft a decision theory onto the components that simulate the attentional and perceptual aspects of the task. It is this flexibility that ultimately makes our models successful.

A Random Walk Model of Search

Random walks are typically applied to study choice decision for a single instance or element. In multiple-target visual search, there are as many as four elements, and this requires that we elaborate the basic model accordingly. We did this by introducing multiple independent walkers, one for each of the *n* elements present in a stimulus display (for related models, see Palmer & McLean, 1995; Ward & McClelland, 1989). In this scheme, any single walk evolves stochastically over time through the formation of a running sum based on repeatedly sampling random deviates (usually Gaussian, though not required). The walks drift in what is formally called a Brownian motion.

In the search model, we interpret the random walk to represent a record of the observer's accumulating evidence about an individual element's identity. Walks corresponding to target elements are formed by summing deviates from a distribution with positive mean and on average move toward a positive target criterion. Similarly, distractor walks are formed by summing deviates from a distribution with negative mean, so that, on average, they tend to drift toward the negative distractor criterion (typically, the means of the target and distractor increment distributions are symmetrically placed about zero, and both distributions have identical variance, though this is by no means required). The overlap of target and distractor distributions may be manipulated to simulate the variation in underlying evidence confusability.

In the simplest form of the parallel model of MTS, *n* random walks drift simultaneously between the two decision criteria. A target-present decision is made when the first walk crosses the positive criterion; a target-absent decision is made when all walks have crossed a negative criterion. Target-present reaction time in the model is given by the number of sampled deviates required for a given walk to reach a response criterion. Errors occur when a walk is absorbed at the wrong response boundary. For example, a false alarm is recorded when one of *n* distractor walks reaches the target-present boundary in error. Schematic details of the algorithm are shown in Figure 10. In the figure, a single target walk and a single distractor walk are illustrated (plotted in black and gray, respectively), along with representations of the underlying increment distributions that give rise to each accumulation process (note that the moments of these distributions are exaggerated in the figure). *S* is the mean displacement of the evidence samples (i.e., the average step size in a random walk) and is what distinguishes targets from distractors in the sampling algorithm. *V* is the common variance of the evidence samples. *T* is the distance of the absorbing boundaries.[5]

Asymmetry in Target-Absent Response

Pilot simulations of the random walk model revealed that the basic architecture is inadequate to deal with certain aspects of response asymmetry. In the first place, people are generally slower to make negative judgments (Baddeley & Hitch, 1974; Clark & Chase, 1972; Kosslyn, 1975; Lewis & Anderson, 1976; Sternberg, 1969; Treisman & Gormican, 1988), and this feature cannot arise if the absorbing boundaries and the evidence distributions are both symmetric about the walk origin. One has to be displaced, and we displaced the distractor boundary so that it is slightly more distal by a factor *D*. *D* is always greater than one in the simulations. The effect of *D* is to delay target-absent decisions by requiring the random walk to migrate further from the origin. This delay has a concomitant effect on error: fewer misses and (a few) more false alarms.

There may also be aspects of decision making that inhibit negative responses that may have nothing to do with the gathering of evidence, and these are not captured by the model in any sense. The evidence that nonvisual and nonattentional processes are active in search is that we on occasion found that RT for signaling the presence of a target at Set Size 1 would sometimes be much less than the RT for signaling the absence of a target at Set Size 1 even though the error rates in the two cases were virtually identical. Where large RT differences for a single object are so large and there is no speed–accuracy trade-off occurring, a secondary process of inhibition is implicated. This model does not attempt to characterize such secondary processes, and when there are large disparities in RT at Set Size 1 with no attending difference in error, we use the difference in time to estimate an overall inhibition cost. This cost is then subtracted from all of the target-absent RTs in models of that experiment to remove the influence of these out-of-model factors.

[5] *T* represents the amount of evidence necessary to decide an element's identity and thus sets the distance of the target-present and target-absent response criteria from the walk origin (the target-absent criterion is by default equal to the target-present criterion but of opposite sign; the walk origin is set to zero). A target-present response is initiated as soon as one of the *n* independent walks reaches the boundary set by *T*; a target-absent response is initiated as soon as all of the *n* walks reach the oppositely signed distractor bound. Because of simple scaling relationships among *T* and the mean (*S*) and variability (*V*) of the increment distributions, one of these parameters can be fixed without loss of generality. To see this, note that *T* is really just a distance that can be expressed as a linear combination of *S* and *V* (i.e., $T = ZV + S$, where *Z* is a normalized distance). For all the modeling work reported here, we have chosen to fix *T* at 20 and to let *S* and *V* enter the model as free parameters. Because of this, the exact value of *T* has no real psychological meaning in and of itself—it is set simply to produce walk times and error rates commensurate to the fitting of data.

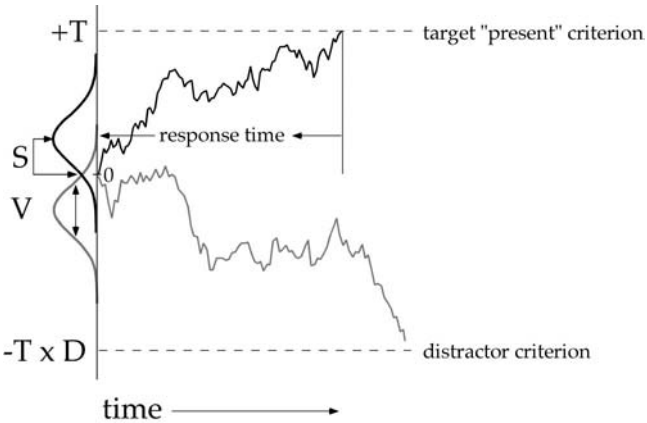

Figure 10. Schematic representation of the core random walk generator used in both the parallel and serial models of multiple-target search. The parameters *S* and *V* correspond to the mean and standard deviation of the increments accumulated in the random walks. The target-present and target-absent criteria are set by the constant *T* and the asymmetry parameter *D*. A single example walk based on increments sampled from the target distribution is shown in black; an example walk for the distractor case is shown in gray. The corresponding Gaussian increment distributions on the left have been exaggerated for illustrative purposes.

A Parallel Model of Multiple-Target Search

The model described here is intended to simulate the search process from first principles. The model acts as an observer in the sense that it makes decisions about stimuli. The stimuli in the model are isomorphs to the stimuli that humans evaluate. So, in parallel search, the model acts upon all elements in the display simultaneously. This is represented in the language of random walks by simply allowing as many random walks as there are elements in the display—recognizing, of course, that this display is nothing more than a set of instructions that tells the model which evidence distributions (target or distractor) to sample from. The walks corresponding to different elements accumulate independently but on the same clock cycle—they all take steps together. A schematic representation of this architecture is shown in Figure 11. In this figure, each random walk is linked to a single element (depicted by grayscale), and the walks all move in the same decision space—the absorbing boundaries are not indexed to the individual random walks.

Set-Size Costs and Capacity Limitation in Parallel Models

Capacity limitation is a construct associated with the spreading of a finite resource over multiple tasks. In the case of multiple random walks, capacity limitation is represented as a slowing down in the rate of accumulated evidence. This slowing is a function of set size, and the rate of evidence accumulation must be strictly decreasing with increasing set size. There are an indefinite number of ways of realizing this mathematical constraint, and the procedure we have chosen is simply to reduce the step size as a function of set size. Specifically, the step size obtained by sampling the evidence distribution is multiplied by $n^{-\gamma}$, where n is the set size and γ is a free parameter that sets the cost of divided attention.[6] When γ is zero, there are no set-size costs, and this means that RT and error will be invariant with display numerosity. In contrast, when γ is close to one, the step size is reduced by the factor $1/n$, and this means that model RT will rise dramatically with display numerosity.

We estimated γ by directly fitting models to data. It is small (on the order of .2) in feature search and as large as .8 in the most difficult configuration searches. This particular manner of enforcing set-size costs increases RT, but it also decreases error rate as the variability of the evidence distribution is also scaled. The qualitative effect of γ on the rate of evidence accumulation is depicted in Figure 11. The random walks in the upper and lower panels differ only in terms of the size of γ: .2 in the upper panel, .8 in the lower.

Relaxation of Decision Criteria in a Parallel Model

The observation that people are actually faster in making target-absent judgments at larger set size mandates that the simulation have some flexibility in its decision protocols. People are faster because they are less accurate, and they choose to suffer the loss in accuracy because targets are in fact rare at large set size and time is indeed of the essence. We required a decision mechanism in the random walk procedure that would allow early absorption because this is how RT is reduced. We introduce here the notion of preponderance of the evidence as a vehicle for early termination. Specifically, we added a second absorption boundary for distractors that has the following logic: If all walks have crossed (i.e., are more negative than) this more proximal boundary, then as there is no early evidence for a target, there probably is no target present—so, terminate the trial. Although this strategy increases the relative frequency that single targets are liable to be missed at larger set size, it has little overall effect on the total error rate.

Preponderance of the evidence is instantiated in the algorithm through a parameter *C*. The relaxed distractor criterion is placed

[6] We reduce the step size by multiplying each evidence sample by a scale factor ($n^{-\gamma}$) that is smaller than one. This is equivalent to reducing both the mean (*S*) and standard deviation (*V*) of the walk increments identically as a function of set size. There are several reasons we chose this implementation of capacity limitation. First, it seemed intuitive to conceive of attention as having a multiplicative influence on stimulus information (Palmer, 1989), much as increasing resistivity has a multiplicative influence on current flow. Second, we chose this type of scaling because it preserves the intrinsic discriminability of target and distractor increments, the aim being to incorporate attentional effects independent of the underlying evidence quality. Finally, only this manner of scaling was able to fit the error patterns observed in the MTS test bed (namely, the decreasing false alarms with set size). An alternative to multiplicative attenuation is to scale the mean and standard deviation of the increments independently. A common approach in this vein is to scale the standard deviation by the square root of the factor that is used to scale the mean (Ward & McClelland, 1989). This kind of scaling is related to sample-size models of attention and leads to decreases in evidence discriminability with set size (cf. McLean, 1999; Palmer et al., 1993; Verghese & Nakayama, 1994). Our parallel model of search does not incorporate such scaling because it quite generally leads to the prediction that false alarms should increase with set size—this is a pattern that never occurs in the use of the MTS method.

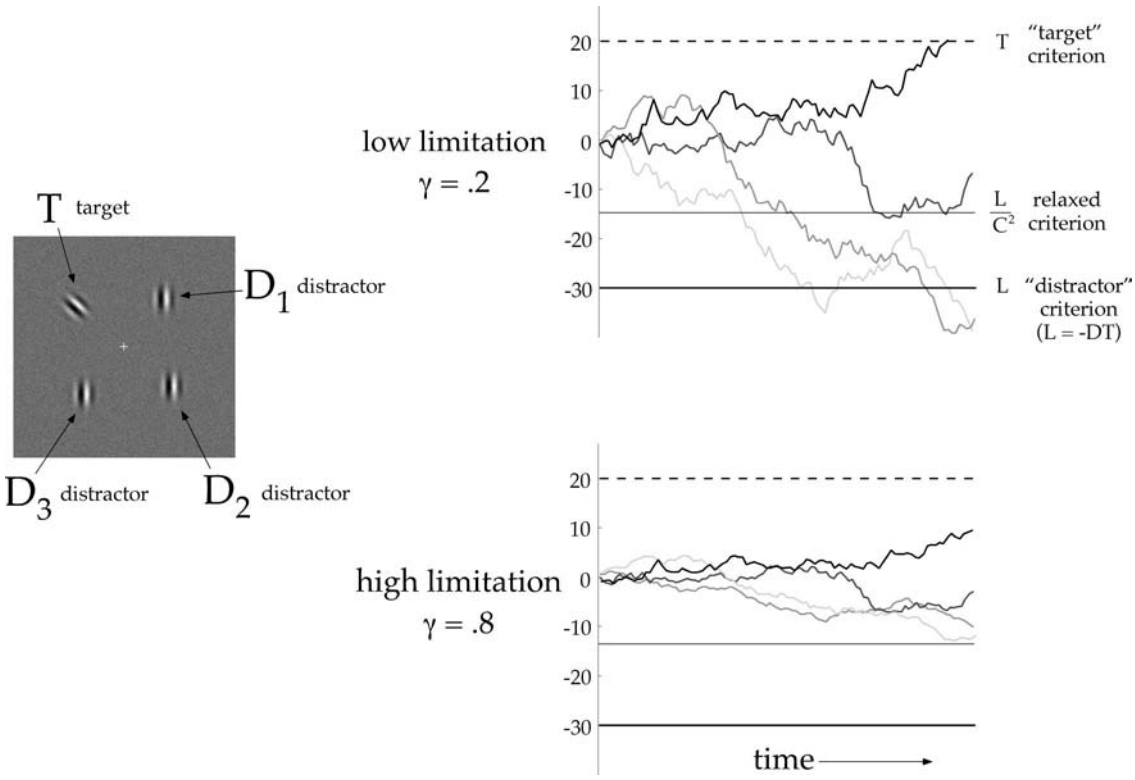

Figure 11. Elaboration of the random walk generator for use in the parallel model of multiple-target search. The various random walks denote the four accumulation processes that correspond to each of the elements in a Set Size 4, single-target display. One walk is based on target increments (black); the remaining walks are based on distractor increments (shades of gray). In the parallel model, the walks diffuse independently and accumulate evidence simultaneously. The attentional limitation parameter γ attenuates the drift rate of each random walk as a function of set size. The parameter *C* controls a secondary joint walk criterion that can trigger a target-absent response depending on the joint state of all *n* walks at a given moment in time.

at a distance $-L/C$ at Set Size 2 and at L/C^2 at Set Size 4, where L is the boundary for distractor absorptions when there is only one walk ($-DT$). C is adjusted by fit to the overall RT and error patterns in each study. This parameter, as expected, played a major role in explaining data. In every case, C was greater than two, meaning, for example, that at Set Size 2, people require about half the evidence from any one element that they would need to identify a single element in isolation.

Underlying the notion of preponderance of the evidence are two facts about the design: Targets are increasingly rare at large set size, and many target-present displays have multiple targets. These facts seem to have been at least implicitly understood by our participants and are presumably why people choose to evaluate displays on the basis of early perceptual returns. Why perseverate in a search that has not produced any target evidence? It is possible to calculate exactly the posterior probability that a target is present given the joint state of *n* walks. This probability allows a boundary to be defined, say, where the probability is .95, so that the error rate is held at .05 for all target-present displays (see Appendix A). We have found that preponderance of the evidence acts as a heuristic that closely approximates the optimal Bayesian boundary placement.

A Serial Model of Multiple-Target Search

There is a long tradition in the memory and visual search literatures of contrasting predictions of parallel and serial models. In virtually every case, the kinds of serial models that have been proposed (and occasionally explicitly evaluated) build on the idealized conception of a serial process in which elements are analyzed independently in sequence, randomly, and without replacement (see Horowitz & Wolfe, 1998, for a memoryless account of serial search). In these kinds of models, each element is thought to receive an identical analysis using a set of invariant response criteria and a processing rate that remains fixed throughout the sequence of identifications. Let us refer to such a serial model simply as a *fixed criterion model.* That it uses fixed criteria leads to two immediate consequences in the domain in which it has been evaluated, that is, STS: Target-present RTs should rise in proportion to set size, and target-absent RTs should increase at twice this rate. This logic is the origin of the expected 2:1 slope ratio that specifies the so-called serial search in single-target methods.

Although this conception of seriality is ingrained in the search literature, it must be recognized that no rational person would search

in this manner unless there were extreme penalties or rewards imposed. Imagine a 12-element display in a design where half the stimuli at each set size have one target and half have no targets. Would it be rational or even reasonable to exhaustively search this display? It is much more likely that a given display has no targets than that the target will be encountered among the final few uninspected elements. Animals that behaved this way looking for food would starve (see Alexander, 1996).

The fixed criterion search model is a poor candidate for explaining any of the MTS data. The consequences of fixed criteria in the multiple-target methodology are easily worked out and are displayed in Figure 12 (left panel). In the right panel of the figure, we show the RT and error-rate data obtained for a typical search based on rotation direction (Thornton & Gilden, 2001), a task that is as difficult as anything we have encountered. It is evident that the fixed criterion search model makes predictions that look nothing like these data, and these data are our best candidates for a serial process. The miss rates, false-alarm rates, and target-absent RTs are in complete disagreement in terms of their global shapes. The fixed criterion model requires flat miss rates and increasing false-alarm rates with set size. This is simply not seen in any of the data.

The problem with this serial model is not that it is serial but that the decision structure is irrational. Computational implementations of this model typically fail to explain data (e.g., Eckstein, 1998; McElree & Carrasco, 1999), and the default model fails to explain our data as well. It is clear that decisions about processing style should not be based on what is essentially a straw man of a serial process and that we must develop the serial process so that it is able to effect rational speed–accuracy trade-offs. This is exactly what the fixed criterion serial model cannot do, and it is why the model is virtually useless. It is basic to any decision procedure that the validity of the distinction is supported only to the extent that the two competing alternatives are equally plausible and similarly articulated so that each has some chance of fitting data. Here, we require a good model of the serial process.

We have developed a serial architecture that is psychologically motivated and sufficiently flexible so that it may compete with the limited-capacity parallel models. Like its parallel counterpart, the serial model of MTS inherits the three core parameters of the random walk generator (*S*, *V*, and *D*). What makes the serial model serial is that it generates walks one at a time. This implies that in a Set Size *n* condition, the model will make *n* successive decisions based on *n* random walks—each walk must terminate in either a target or distractor identification prior to initiating the next accumulation process. If one of the random walks is absorbed at the target-present criterion *T* (prior to absorption at the distractor criterion), a "target-present" response is generated, and no further identifications are carried out. Provided that no random walk crosses the target-present boundary, the sequence of analyses will continue until all *n* elements have been categorized.

This kind of serial decision structure is generic to serial models but is insufficient for our purposes. A useful serial model must also be endowed with flexible criteria if it is to fit any of the data in MTS. Realistic serial models require that the decision criteria be sensitive to both set size and evidence accumulated during the trial. These adjustments basically allow the serial model to save time by not perseverating in fruitless searches for rare targets.

Relaxation of Decision Criteria in a Serial Model

The motivations for introducing the notion of preponderance of the evidence and the parameter *C* in the parallel model apply here as well. Although the preponderance of current evidence is not an issue in serial access to information, we must allow

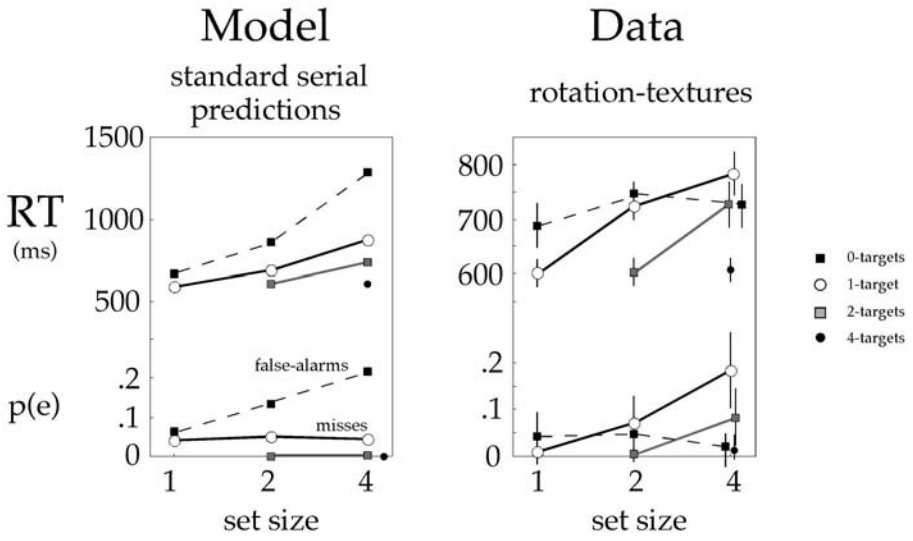

Figure 12. Comparison of the multiple-target search predictions of a standard serial model having invariant criteria and exhaustive processing (left panel) with actual data from the rotation-textures experiment (right panel). Error bars denote standard error of the mean. RT = response time.

criteria to be set-size dependent if the model is to not waste time on the multiplicity of displays that in fact contain no targets.[7] In practice, we implemented this kind of rationality in the serial random walk model via the parameter β, which moves the base distractor criterion (*–DT*) closer to the origin as set size increases. A schematic representation of the effects of β on the distractor criterion is shown in Figure 13. Note that β has its influence only on the distractor criterion and that the target response criterion remains fixed at *T* in all cases. This manner of asymmetric β-scaling was adopted because it alone generates patterns of RT and error consistent with the observed data.

Positive values of β generate several clear effects on the predicted patterns of RT and error in a serial model. For $\beta > 0$, the distractor criterion is moved relatively closer to the origin, and thus, distractor elements are identified more quickly—but with different consequences for the various kinds of error. False alarms are reduced, whereas miss rates increase (both consequences of easier access to the distractor boundary). Such set-size effects are in fact common in the most difficult searches (Class C). However, if we want a serious serial competitor to the parallel limited-capacity model of MTS, there is an additional source of decisional flexibility to be implemented.

Relaxation During the Trial

The second component of the serial model is its decision parameter T_Z. We included T_Z to allow the simulated observer to incrementally relax its distractor criterion in situations where it is uncovering nothing but distractors. A flexible decision heuristic such as this allows the observer the freedom to conduct a cursory inspection of remaining items so as to end the trial in a timely manner. The additional error that is accrued is more than outweighed by the time saved in avoiding fruitless searches of the myriad number of target-absent displays (recall that at each set size, the number of target-absent displays must balance all of the ways that targets may appear; see Figure 2). This is exactly the freedom that a serial observer must have to rationally conserve a valuable resource such as time. In this way, only a fraction of the elements in a display is fully analyzed, and searches are effectively abandoned if the elapsed time approaches some fixed deadline. This notion is similar to many of the guessing rules used by high-threshold models of search (see Chun & Wolfe, 1996; Palmer et al., 2000).

A schematic representation of how T_Z is implemented in our random walk framework is shown in Figure 14. Here, we depict a hypothetical analysis of a display containing four distractors. The leftmost random walk represents the accumulation of evidence of the first element to be inspected. This walk terminates at C_1, the base distractor criterion set by $-DT$ and β. The time required to achieve a categorization of the first element as a distractor is denoted t_1. Prior to analysis of the second of the four elements, the criterion C_1 is relaxed by an amount equal to the fraction of T_Z remaining after the first analysis (i.e., $1 - t_1/T_Z$). This criterion, C_2, sets the boundary for the next distractor decision. This procedure successively repeats until all four elements have been inspected, a target has been found, or a time T_Z has elapsed. Once a time T_Z has elapsed, the distractor criterion is at the origin, and no further trial time is allowed to accrue (decisions are based on the sign of the next random deviate).

Summary of the Models

We have developed two models of MTS based on noisy accumulation of evidence regarding element identity. One model adopts a parallel, limited-capacity architecture, the other a serial architecture. The models both share an identical core random walk generator parameterized by the variables *S*, *V*, and *D*. These parameters determine in part the single-element decision times and error rates. The parallel model elaborates on the basic discriminator by introducing a capacity-limitation parameter (γ) that multiplicatively attenuates the magnitude of evidence available to decision as a function of set size. The model also has a parameter that allows decision criteria to relax with set size (*C*). The serial model has two unique parameters that serve to relax decision criteria as a function of both set size (β) and accumulated processing time (T_Z). Like the parameter *C* in the parallel model, these parameters allow the serial model to incorporate a rationallike strategy that is responsive to changing priors on element identity during search. The formal functions of the full set of shared and model-specific parameters are summarized in Figure 15, along with the principle psychological motivations supporting the inclusion of each parameter.

Heuristic Guide to the Serial–Parallel Decision Problem

The most straightforward way to solve the parallel–serial classification problem is through brute-force simulation with model selection based upon goodness of fit.[8] As the complete data set comprises nine RTs and nine error rates, the complexity of the data patterns mandates a formal procedure. However, before we present algorithms for model selection, it is instructive to introduce the key signatures in data that distinguish the serial and parallel processes. We present these signatures as heuristics in deciding the serial–parallel issue. In spirit, these heuristics are similar to the rule that if target-present RT in-

[7] Criterion relaxation with set size is rational regardless of processing style. The reason is simply that targets become increasingly rare at larger set size, and if time is a commodity, it is not advantageous to waste it searching for targets that are probably not there. We have computed the extent to which distractor elements outnumber target elements in MTS on the basis of the probabilities of encountering specific stimulus types and a tally of the number of target and distractor elements present in each type. For Set Sizes 2 and 4, the probabilities of encountering a distractor are 5/8 and 17/24, respectively. The implication of this imbalance in the priors, as far as decision is concerned, is that for set sizes greater than one, it is almost twice as likely that any given element currently under inspection will be a distractor.

[8] Because the two computational models we have created require complex nonlinear decision structures, they are not amenable to simple analytic decomposition (Laming, 1968). To evaluate these kinds of models, we rely instead on brute-force Monte Carlo simulations to generate predictions. Running the model in an explicit simulation is the only way to determine exactly what the model will do in any given situation, though, in simple limiting cases, we do check that the simulations give mathematically correct answers using analytic formulations based on either continuous diffusion to a boundary (Ratcliff, 1978) or the discrete random walk (Karlin & Taylor, 1975; Smith & Mewhort, 1998). What numerical simulation lacks in elegance, it far makes up for in terms of flexibility—by forgoing the world of analytic model fitting, it is possible to explore the variety of complex decision structures embodied in the parallel and serial models of MTS we contemplate.

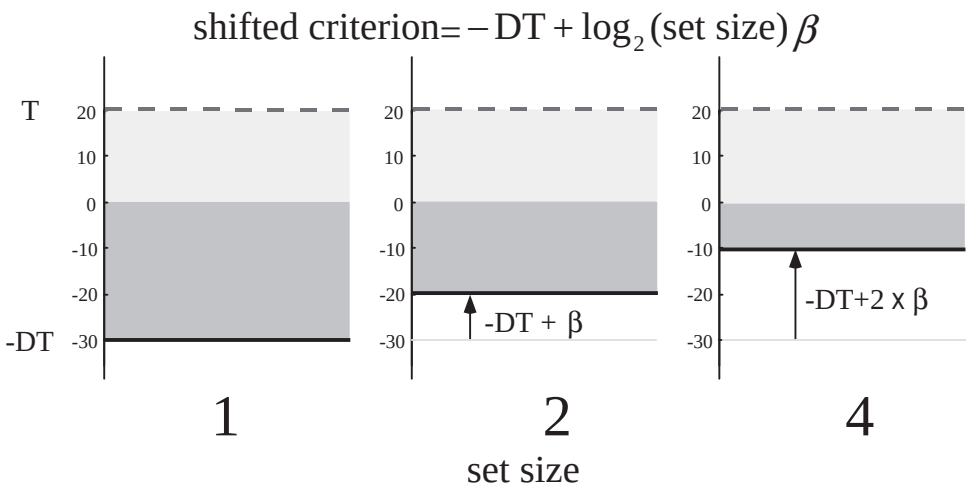

Figure 13. Schematic representation of the effect of the serial model's set-size relaxation parameter β. At set sizes of 2 and 4, β scales the target-absent criterion (prior to analysis) so that it is closer to the origin. $-DT$ = base distractor criterion; T = target-response criterion.

creases with set size, then serial, which guided early research in this field. Our heuristics are more subtle and more numerous because of the fact that it is necessary to consider both speed and accuracy in deciding the kind of search process that is operating. We have found that going through the following rules, observations, and constraints is inevitable in conceptually navigating the enormous complexity of search data.

Heuristic for Both Serial and Parallel Models

The Probability of a Miss

All models with fixed decision criteria tend to produce shallow or flat miss rates on single-target trials.

Observation

When there are single targets among a variable number of distractors, the miss rate (responding "target-absent" on target-present trials) increases with set size throughout all data sets.

Heuristic 1

Serial and parallel models must allow distractor boundary movement that is contingent upon set size. As the distractor boundary moves inward, the probability that any random walk will be absorbed on the distractor boundary increases. In this way, the random walks that represent targets tend to be missed. The benefit of making misses on single-target trials is that observers do not become bogged down on target-absent trials while they make sure that each distractor actually is a distractor.

Heuristics Related to the Serial Model

Probability Summation

A serial search with a fixed target-present criterion must produce false alarms (responding "target" on target-absent trials) that increase with set size. The probability of an error (false alarm) must increase when there are more items to be rejected.

Consequence

If the false-alarm rate at Set Size 1 (fa_1) is large, say, 5%, then the false-alarm rate at Set Size 4 will be about 20% in a serial model with fixed target-present criterion.

Observation

There are no increasing false-alarm trends in the data.

Heuristic 2

A serial model with a fixed target-present criterion must be constructed so that fa_1 is near zero (such models have traditionally been referred to as *high-threshold models*; Palmer et al., 2000). If the serial model is to fit data with fa_1 on the order of 5% or larger, then it cannot have fixed decision criteria: The model must then allow the target-present criterion to move outward with set size to suppress the effects of probability summation.

Speed–Accuracy Trade-Off

Serial models that allow the target-present criterion to move outward with set size produce steep costs in reaction time. Even the pure-target RTs will increase dramatically with set size when the criterion is moved so as to suppress the effects of probability summation.

Observation

There are no steep pure-target functions in the data. Where fa_1 is 5% or greater, pure-target functions are generally flat or decrease with set size.

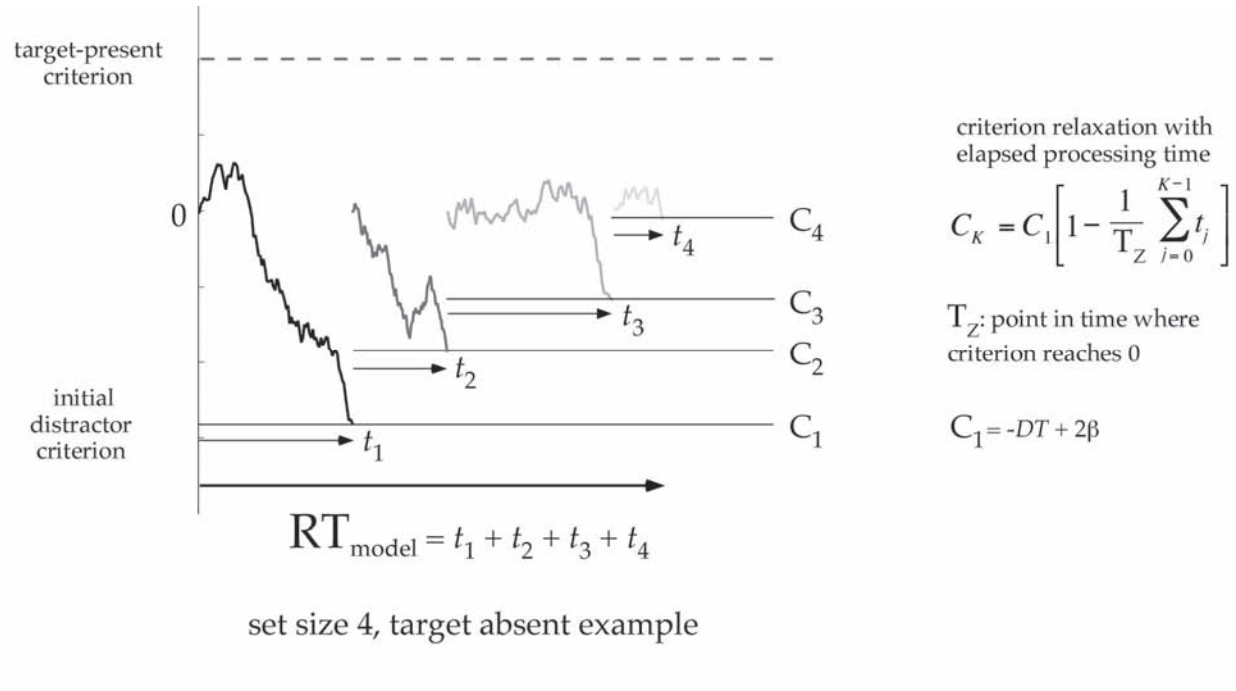

Figure 14. Schematic representation of the effect of the serial model parameter T_Z for a Set Size 4 display. With each subsequent classification of an element as a distractor, T_Z proportionally relaxes the target-absent criterion (within a sequence of element analyses) as a function of the total elapsed processing time. RT = response time; $-DT$ = base distractor criterion; C_k = decision criterion for element k; t_k = time to make decision on element k.

Heuristic 3

The only data sets that are candidates to be fit by a serial model will have small fa_1.

Heuristic Related to the Parallel Model

Redundancy Gains

The distribution of fast finishers will always be speeded relative to the finishing time of a single process when there is little cost for divided attention (Raab, 1962). In this way, parallel processes with small γ typically show statistical race gains. A parallel search with flat or rising pure-target function must have γ on the order of 1 (this is the largest γ can be in this model when attentional resources are divided exactly by the set size).

Divided Attention Lowers Error

If all other variables are held constant, accuracy increases monotonically with increasing γ. Divided attention is implemented through smaller step size. Smaller step size means that more steps (n) will be required for any given walk to be absorbed at any boundary, and error rates grow relatively as $1/\sqrt{n}$. In particular, it is difficult for a high-γ model to produce large miss rates.

Speed–Accuracy Trade-Off

Miss rates constrain the slopes of the target-absent RTs with set size. If a distractor boundary is moved inward to force the parallel model to accept larger miss rates, the model will necessarily produce low set-size costs in the target-absent RTs.

Heuristic 4

The parallel model is ill suited to fit data in which both the miss rate and target-absent RT increase with set size.

Summary

There is a class of data that has small fa_1, flat false-alarm functions, rapidly increasing miss-rate functions, and increasing target-absent RT functions. This one class is best fit by a serial process. The remaining data are characterized by decreasing false alarms and trade-offs between miss rates and speed on target-absent trials. These data are fit by the parallel model.

Deciding Serial or Parallel by Goodness of Fit

The central goal of our work is the development of a formal procedure for deciding how element analysis is scheduled in visual search. As has been previously noted, these models have many parts and are addressing complex patterns in data. For their mechanics to be comprehensible, it is necessary to describe the procedure in some detail. All the insights and generalizations that we draw from this work ultimately come from the certainty that may be attained by using a rote algorithm.

Normalization of the Models

The serial–parallel problem in MTS data is decided by assessing the proximity of models to all the available data. The issue of proximity is not entirely straightforward as the model RT and human RT are offset by aspects of response that go into perceptual encoding and keypress response (response mapping, execution, etc.). To bring the two systems of RT into alignment, we normalized the models so that the average of the pure-target RTs is the same for both model and data. Although there are certainly other ways of taking into account elapsed time produced by nondecisional components of the task, we have found that this procedure constrains the chi-square fit to capture the most important trends (pure targets, target absent, and single target). Error rates in model and data are also slightly out of alignment because of mistakes in response mapping, various forms of distraction, pushing the wrong finger, and so on. Misalignments in error are generally caused by sources of confusion or uncertainty that are outside of the purview of the diffusion framework. The best estimate we have of extraneous error is the error rate in the four-target condition. Whenever this error is the smallest over the entire design, we align the error rates in model and data by subtracting it from all cells. Otherwise, no alignment is attempted. Aligning RT and error between model and data has no effect on the relative patterns, but it is required for estimates of goodness of fit.

Reconciling Reaction Time and Error Residuals

In the fitting of models to data, we have had to reckon with the obvious problem that error rates and reaction times are measured on different continua. This problem has rarely been faced in the search literature, or in any psychological literature for that matter, insofar as models tend to be applied to RT or error—but not to both, as we are attempting to do. One strategy that has been adopted to circumvent this apples-and-oranges problem is to apply an arbitrary multiplier to RT residuals so that they can be meaningfully combined with error residuals (Maddox & Ashby, 1996; Van Zandt et al., 2000). We prefer to use the natural metric provided by chi-squares, using intrinsic variability to measure the distance between model and data. In this method, RT chi-square is added to error chi-square to obtain a total chi-square for the entire data set. As usual, the model with the minimum chi-square is selected as the best description of the data in question. This process not only provides a solution to the serial–parallel classification problem but also generates meaningful parameter estimates that characterize both search difficulty and how people adjust their decision criteria.

	parameter		principle
Random walk generator	S	mean of increment distributions	controls the rate of evidence accumulation
	V	increment variability	accumulation process is stochastic
	D	target-distractor criterion asymmetry	single target RTs are faster and more accurate
Parallel model	γ	attentional limitation	attention multiplicatively scales evidence
	C	criterial relaxation with set size	decisions can be made using weak, but consistent lines of evidence
Serial model	β	set size-dependent bias	distractor elements are more likely to be encountered at large set sizes
	T_Z	criterial relaxation with processing time	only the first few elements receive a detailed analysis

Figure 15. Summary of model parameters. RT = response time.

Specifications of the Decision Algorithm

The serial–parallel decision algorithm is spelled out briefly here and in further detail in Appendix A.

1. Find the parallel process that minimizes the chi-square deviation between model and data (do this for 100 randomly resampled pseudoexperiments).

2. Find the serial process that minimizes the chi-square deviation between model and data using the same resampled data as in Step 1.

3. For each resampled pseudoexperiment for each task, compute the difference in chi-square between the best parallel and serial models ($\Delta\chi^2$). Average the differences over the 100 pseudoexperiments, and estimate the 95% confidence limits. By virtue of the independence of data resampling, $\Delta\chi^2$ is proportional to the log-likelihood ratio.

4. The best model for each task is given by the sign of $\Delta\chi^2$ (i.e., the log-likelihood ratio). As we have an estimate of the variability of $\Delta\chi^2$, we can assess the reliability of each decision.

Results of the Decision Procedure on the Test Bed of Data

The results of applying this algorithm to each of the 29 tasks in the ensemble are shown in Figure 16, where we have plotted the log-likelihood ratio and the associated 95% confidence limits for all 29 tasks. Parallel tasks are on the left (negative logarithm), and serial tasks are on the right (positive logarithm). The ensemble has been further ordered from top to bottom using the standard measure of search efficiency (i.e., the single-target RT slope) computed from the data. The model-based estimates of attentional limitation via the parameter γ have also been inset for those data sets fit best by the parallel model.

For the most part, there is always one process that fits the data in each task quite well. In Appendix B, we have plotted every data set in the entire ensemble, together with the best fitting model. The majority of tasks (21) in the ensemble appear to be mediated by a parallel process. These tasks include feature conjunction, letter identity, shape identity, and relative-position judgment. The remaining 8 tasks—rotations and mirror inversions of luminance polarity (shading-LR, phase) and arbitrary coalitions of circles and plusses—are serial. It must be recognized that this ensemble is not representative of either natural search or laboratory search tasks. The probability of selecting a serial task from the literature is not

8/29 but much smaller. As the construction of the task ensemble was partly guided by our interest in difficult search, the membership is biased toward serial processes. Nevertheless, demonstrating the existence of a class of serial processes is one of our principal findings in that there has been a growing tacit assumption in the literature that all visual search is conducted in parallel with more or less capacity limitation (Palmer, 1995; Palmer et al., 1993; Pashler, 1998; Wolfe, 1998a).

That we have been able to fit well over 20 data sets with a pair of fairly simple models suggests that the basic mechanisms of decision and attention are replicated within the model architectures. To this extent, we have captured the strategies and mental set that are involved in looking for something. However, there is one task that seems to involve different strategies, and neither our serial nor our parallel model provides an adequate fit to both RT and error; this task is orientation search. This failure occurred primarily because observers had a relatively difficult time responding correctly (error rates almost 10%) to the identity of distractors and targets when presented in homogeneous fields, that is, on pure-target trials and on target-absent trials. Category confusion could have occurred if observers did not in fact regard the targets and distractors as being in different categories. In this sense, an oriented Gabor (or line) is just one single object regardless of its orientation, and it may have been difficult for our observers to assign categories to the particular angle at which it was viewed. The orientation data are consis-

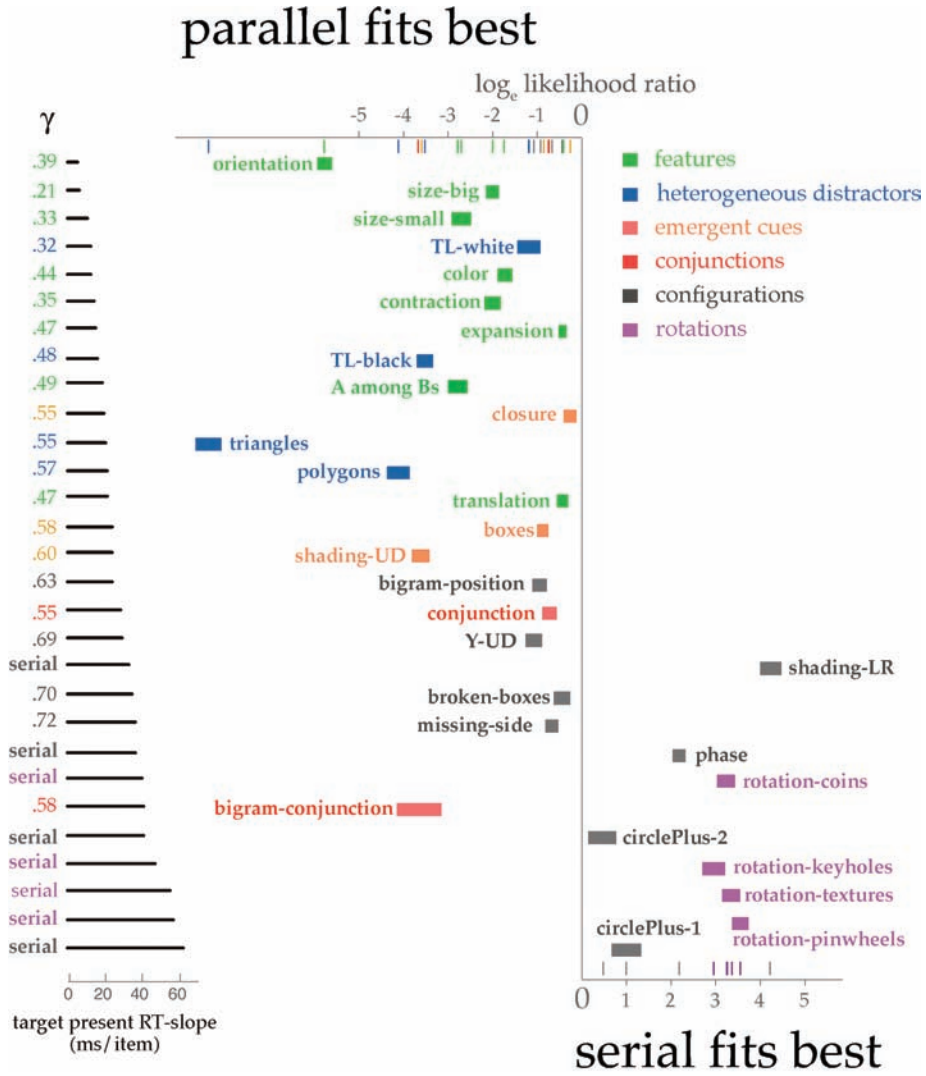

Figure 16. Competitive modeling of the multiple-target search (MTS) ensemble based on fits of the parallel and serial models to each data set. The abscissa plots the average log-likelihood ratio (with 95% confidence limits) for all 29 data sets. Tasks falling to the left of 0 were fit better by the parallel model of MTS; those falling to the right were fit better by the serial model of MTS. The entire test bed is ordered using the standard measure of efficiency (i.e., the single-target response time [RT] slope across set size). The value of the capacity-limitation parameter γ is included for the 21 tasks in the test bed best fit by the parallel model.

tent with this hypothesis, and both RT and error appear to be mediated by a search strategy that is aided by heterogeneity. Heterogeneity implies that a target must be present, whereas homogeneity has no entailment—hence, the relatively large error rates. The unique and odd error patterns aside, the RT patterns are indicative of a high-capacity (low-γ) parallel process (as illustrated in Figure 9A), and the best fitting model is parallel. We replicated these results on several separate occasions using oriented lines, oriented Gabors, two sets of absolute angles, and two sets of eight observers. Indeed, this enterprise would be compromised if we were to conclude that orientation search was serial.

Model Parameters

The best model provides not only a protocol (serial or parallel) but also a parametric description of the search process. For example, we know how observers evaluate the preponderance of multiple lines of subthreshold evidence (*C*), how effectively they are able to divide attention across elements (γ), and how they manage time in slow serial search (β and T_Z). γ is of particular interest because it provides an effective ordering of all parallel search processes.[9] γ is distinguished from simple data features, such as the slope of the target-present RT with set size, in that it literally incorporates aspects of all 18 degrees of freedom. Still, the ordering implied by the magnitude of γ is not meaningfully different from a standard slope ordering (shown in Figure 16), though there are several tasks where γ and the standard measure of efficiency disagree (e.g., in both conjunction tasks [red] and for orientation). We also find that all parallel search tasks lead to a common form of criterion movement, and in Appendix A, we show that this movement is rational in the context of multiple-target parallel search (see the Appendix A section entitled Rationality of Criterion Relaxation and the Parameter *C*). Similarly, we find a common level of criterion movement in serial search data that is also consistent with rational decision making—observers are able to sacrifice accuracy for speed in a principled way by abandoning those searches that are least likely to yield a target. The detailed model parameters that summarize attentional limitation and boundary movement in all 21 parallel tasks and all 8 serial tasks are given in Appendix C.

The Natural Kinds of Search

Previous research has also found that search data divide into roughly three categories, which we have called the A, B, and C classes. Much of the history of this field can be traced in terms of how these classes have been regarded. Originally, Classes B and C were thought to be serial (Treisman, 1988; Treisman & Gelade, 1980), and certain aspects of the RT patterns certainly do look serial. With the advent and popularization of the concept of capacity limitation (Townsend, 1972, 1974, 1990), it became clear that there was no reason to distinguish any of the classes on the basis of STS data. As a consequence, all three classes are now regarded as residing on a continuum of capacity—or, in more modern language, efficiency (Wolfe, 1998a).

The work presented here resolves some of these issues. First, the universe of search data cannot be placed onto a continuum of efficiency. There is evidence for both serial and parallel search tasks, and these operate via different protocols. They cannot be smoothly morphed into each other. However, this does not mean that there is a sharp dividing line between serial and parallel search in the observed data. Two well-defined categories may have a fuzzy evidentiary boundary, and such appears to be the case here. For example, the conjunction bigram task generated data that had parallellike false-alarm rates but seriallike speed–accuracy tradeoffs between the miss rates and target-absent RT. Neither the serial nor the parallel model produced satisfying fits. We note that this is a task near the serial–parallel boundary—of all parallel tasks in the ensemble, it produced the greatest single-target slopes.

Perspective

In this article, we have presented a general framework that marries an improved methodology to explicit computational theories of attention and decision. The framework demonstrates that the serial–parallel issue can be resolved in the domain of simple visual search. Admittedly, our conclusions are relative to the two models we have developed, but these models do manage to capture most of the features in both RT and error. Still, there is little doubt that these models are not the final word, and there are surely insights and improvements that we have not incorporated. One potentially fruitful direction is to create hybrid models that have both serial and parallel aspects in their architecture. Wolfe (2003) has suggested such a model wherein the serial and parallel modes are distinct, temporally separated stages. A hybrid model that has obvious appeal is one in which elements within small clusters are analyzed in parallel while clusters themselves are analyzed in serial succession. Such models might be required for understanding search behavior at large set size (say, more than eight elements) when search is known to be hard, even at two or four elements. For example, the classic letter search (rotated Ts among rotated Ls) was found to be parallel at small set size, but there could be a transition to seriality as fine divisions of attention become untenable at large set size. There is, however, always the concern that seriality at large set size may be induced by loss of visual resolution (Geisler & Chou, 1995) and not by the exhaustion of attentional resource.

Our serial model could also be extended to allow for imperfect memory. In large set-size searches, there is evidence that people do not remember what they have discarded (Horowitz & Wolfe, 1998,

[9] The observed variations in γ reported here are not due to task-specific differences in the discriminability of single-target and distractor elements. Some of the most demanding searches (high γ) have stimulus sets that are highly discriminable at the single-element level (Thornton, 2002). Specifically, psychophysical control experiments have verified that even though the missing-side and Y-UD tasks have the largest levels of attentional limitation as measured by γ, they contain some of the most discriminable elements in the MTS ensemble. Of course, manipulations of target–distractor similarity influence our measures of attentional load. When target–distractor stimuli are made highly similar, both RT and errors soar (Duncan & Humphreys, 1989), and this will generally lead to increases in γ. The important point is that all the stimulus sets used here are clearly suprathreshold and thus are near ceiling in terms of single-element discrimination.

2001, 2003; but see Gibson, Li, Skow, Salvagni, & Cooke, 2000; Kristjansson, 2000), and if our model is to be extended to set sizes larger than four, it might be desirable to model item selection in a way that is guided by empirical data. There are also refinements that could be made to our formulation of parallelism. Kim and Cave (1995), for example, showed in a conjunction search that there is nonuniform sampling within the distractor field. Not all features are equally salient, and attention is guided more by color than by shape. In our simulations, we sampled from all elements at the same rate. Parallelism, however, need not imply that all random walks drift at the same rate. Allowing different dimensions to be sampled at different rates would provide a model of parallelism closer to that actually achieved by people. A model able to recognize that objects may vary simultaneously on more than one dimension is considerably more complex than those considered here. Our models are quite simple in that objects are either distractors or targets, and what is being accumulated in a random walk is abstract evidence for the element being one or the other.

What our models do provide is a principled division between serial and parallel processes that is ultimately based on detailed fits to data within a rather large corpus of studies. Almost all of the search data are fit by only one of the two models, and where the models fit, there is external validity. The parallel model fits all of the feature search data (this has to be parallel), the conjunction data (there is some evidence that this is parallel—Eckstein, 1998; Eckstein et al., 2000; McElree & Carrasco, 1999; Palmer et al., 2000), and the emergent cue data (Enns & Rensink, 1990, made a good case for this). The serial model fits only the most difficult search data, those with the largest set-size costs and overall largest latencies. The model parameters also provide a process definition of the notion of efficiency. The set-size cost that we have expressed in terms of γ is effectively a metric of efficiency. Moreover, because γ is estimated using a small number of foveally presented elements, it reveals limitations due solely to attention and is not susceptible to the low-level artifacts that often contaminate estimates of efficiency in large set-size tasks (Geisler & Chou, 1995; Palmer, 1995).

Though we have made substantial progress toward solving the character of attentional limitation in visual search, there are many questions that remain unanswered. From our perspective, the most intriguing of these concerns how aspects of shape determine seriality in search. For example, why are certain configural tasks (e.g., phase, circlePlus) processed serially, whereas other tasks (e.g., missing-side, Y-UD) that seem to be equivalent permit limited-capacity, parallel processing? Are discriminations of rotation direction special within the serial class, or is there a common organizing principle at work that unites them with certain aspects of shape? Within the framework developed here, these questions may now be posed with the certainty that at least one knows that one has a meaningful question and that the assignments serial and parallel can be applied according to a principled procedure.

This procedure is based on simulation for the simple reason that these data are complicated. The history of this field has been to visually inspect the RT data, make simple determinations of slope, check that no obvious speed–accuracy trade-offs are occurring, and then make a decision about process. However, process cannot be reduced to the slope of target-present RT with set size (Treisman & Gelade, 1980). Neither can it be reduced to the slope of the pure-target trial RTs (Townsend, 1990; van der Heijden, 1975). In fact, it cannot be reduced to RT. This field must reckon with process models that handle specific trade-offs between speed and accuracy, and the ones we have invented provide a start in this direction.

References

Alexander, R. M. (1996). *Optima for animals.* Princeton, NJ: Princeton University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

Beck, J. (1966, October 28). Perceptual grouping produced by changes in orientation and shape. *Science, 154,* 538–540.

Bergen, J. R., & Julesz, B. (1983). Rapid discrimination of visual patterns. *IEEE Transactions on Systems, Man, and Cybernetics, 13,* 857–863.

Carrasco, M., & Frieder, K. S. (1997). Cortical magnification neutralized the eccentricity effect in visual search. *Vision Research, 37,* 63–82.

Carrasco, M., McLean, T. L., Katz, S. M., & Frieder, K. S. (1998). Feature asymmetries in visual search: Effects of display duration, target eccentricity, orientation and spatial frequency. *Vision Research, 38,* 347–374.

Carrasco, M., & Yeshurun, Y. (1998). The contribution of covert attention to the set size and eccentricity effects in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 673–692.

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

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3,* 472–517.

Dennis, I., & Evans, J. B. T. (1996). The speed–error trade-off problem in psychometric testing. *British Journal of Psychology, 87,* 105–129.

Duncan, J. (1989). Attention and reading: Wholes and parts in shape recognition—a tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 39–61). Hillsdale, NJ: Erlbaum.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Eckstein, M. P. (1998). The lower efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological Science, 2,* 111–118.

Eckstein, M. P., Beutter, B. R., & Stone, L. S. (2001). Quantifying the performance limits of human saccadic targeting during visual search. *Perception, 30,* 1389–1401.

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & Psychophysics, 62,* 425–451.

Egeth, H. E., & Dagenbach, D. (1991). Parallel versus serial processing in visual search: Further evidence from subadditive effects of visual quality. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 551–560.

Egeth, H. E., & Mordkoff, J. T. (1991). Redundancy gain revisited: Evidence for parallel processing of separable dimensions. In G. R. Lockhead & J. R. Pomerantz (Eds.), *The perception of structure: Essays in honor of Wendell R. Garner* (pp. 131–143). Washington, DC: American Psychological Association.

Enns, J. T., & Rensink, R. A. (1990, February 9). Influence of scene-based properties on visual search. *Science, 247,* 721–723.

Geisler, W. S., & Chou, K. (1995). Separation of low-level and high-level factors in complex tasks: Visual search. *Psychological Review, 102,* 356–378.

Geisler, W. S., Stern, L., Thornton, T., Kuyel, T., & Ghosh, J. (1998). Bayesian ideal-observer analysis of texture segregation: Evidence for structure-based models. *Investigative Ophthalmology and Visual Science Supplement (ARVO), 39,* S649.

Gibson, B. S., Li, L., Skow, E., Salvagni, K., & Cooke, L. (2000). Memory-based tagging of targets during visual search for one versus two identical targets. *Psychological Science, 11,* 324–328.

Gilden, D. L., & Kaiser, M. K. (1992, May). *Kinematic specification of depth and boundary.* Paper presented at the annual meeting of the Association for Research in Vision and Ophthalmology, Sarasota, FL.

Graziano, M. S. A., Anderson, R. A., & Snowden, R. J. (1994). Tuning of MST neurons to spiral motions. *Journal of Neuroscience, 14,* 54–67.

Green, D. M., & Swets, J. A. (1988). *Signal detection theory and psychophysics.* Los Altos, CA: Peninsula Publishing.

Grossberg, S., Mingolla, E., & Ross, W. D. (1994). A neural theory of attentive visual search: Interactions of boundary, surface, spatial, and object representations. *Psychological Review, 101,* 470–489.

Horowitz, T. S., & Wolfe, J. M. (1998, August 6). Visual search has no memory. *Nature, 394,* 575–577.

Horowitz, T. S., & Wolfe, J. M. (2001). Search for multiple targets: Remember the targets, forget the search. *Perception & Psychophysics, 63,* 272–285.

Horowitz, T. S., & Wolfe, J. M. (2003). Memory for rejected distractors in visual search? *Visual Cognition, 10,* 257–298.

Humphreys, G. W., & Muller, H. J. (1993). Search via recursive rejection (SERR): A connectionist model of visual search. *Cognitive Psychology, 25,* 43–110.

Humphreys, G. W., Quinlan, P. T., & Riddoch, M. J. (1989). Grouping processes in visual search: Effects with single- and combined-feature targets. *Journal of Experimental Psychology: General, 118,* 258–279.

Julesz, B., & Hesse, R. I. (1970, January 17). Inability to perceive the direction of rotation movement of line segments. *Nature, 225,* 243–244.

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2005). *Principles of neural science* (4th ed.). New York: McGraw-Hill.

Karlin, S., & Taylor, H. M. (1975). *A first course in stochastic processes* (2nd ed.). New York: Academic Press.

Kim, M.-S., & Cave, K. R. (1995). Spatial attention in visual search for features and feature conjunctions. *Psychological Science, 6,* 376–380.

Kosslyn, S. M. (1975). Information representation in visual images. *Cognitive Psychology, 7,* 341–370.

Kristjansson, A. (2000). In search of remembrance: Evidence for memory in visual search. *Psychological Science, 11,* 328–332.

Laming, D. (1968). *Information theory of choice reaction times.* New York: Academic Press.

Lewis, C. H., & Anderson, J. R. (1976). Interference with real world knowledge. *Cognitive Psychology, 7,* 311–335.

Link, S. W. (1975). The relative judgment theory of two-choice response time. *Journal of Mathematical Psychology, 12,* 114–135.

Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1015–1036.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

Maddox, W. T., & Ashby, F. G. (1996). Perceptual separability, decisional separability, and the identification–speeded classification relationship. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 795–817.

Malik, J., & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America, A: Optics and Image Science, 7,* 923–932.

McElree, B., & Carrasco, M. (1999). The temporal dynamics of visual search: Evidence for parallel processing in feature and conjunction searches. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1517–1539.

McLean, J. E. (1999). *Processing capacity of visual perception and memory encoding.* Unpublished doctoral dissertation, University of Washington.

Meyer, D. E., Irwin, D. E., Osman, A. M., & Kounios, J. (1988). The dynamics of cognition and action: Mental processes inferred from speed–accuracy decomposition. *Psychological Review, 95,* 183–237.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology, 14,* 247–279.

Moore, C. M., Egeth, H., Berglan, L. R., & Luck, S. J. (1996). Are attentional dwell times inconsistent with serial visual search? *Psychonomic Bulletin & Review, 3,* 360–365.

Moore, C. M., Elsinger, C. L., & Lleras, A. (2001). Visual attention and the apprehension of spatial relations: The case of depth. *Perception & Psychophysics, 63,* 595–606.

Mullin, P. A., Egeth, H. E., & Mordkoff, J. T. (1988). Redundant-target detection and processing capacity: The problem of positional preferences. *Perception & Psychophysics, 6,* 607–610.

Najemnik, J., & Geisler, W. S. (2005, March 17). Optimal eye movement strategies in visual search. *Nature, 434,* 387–391.

Nakayama, K., Wang, D., & Kristjansson, A. (2000). Efficient visual search not explained by local feature mechanisms or top-down guidance. *Investigative Ophthalmology and Visual Science Supplement (ARVO), 41,* S579.

Pachella, R. G. (1974). The interpretation of reaction time in information processing research. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41–82). Hillsdale, NJ: Erlbaum.

Palmer, J. (1989, August). *A multiplicative attenuation model of attention.* Paper presented at the annual meeting of the Society for Mathematical Psychology, Irvine, CA.

Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research, 34,* 1703–1721.

Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect. *Current Directions in Psychological Science, 4,* 118–123.

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 108–130.

Palmer, J., & McLean, J. (1995, August). *Imperfect, unlimited-capacity, parallel search yields large set size effects.* Paper presented at the annual meeting of the Society of Mathematical Psychology, Irvine, CA.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40,* 1227–1268.

Pashler, H. (1987). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception & Psychophysics, 41,* 191–201.

Pashler, H. (Ed). (1998). *Attention.* Hove, England: Psychology Press.

Pike, R. (1973). Response latency models for signal detection. *Psychological Review, 80,* 53–68.

Poder, E. (1999). Search for feature and for relative position: Measurement of capacity limitations. *Vision Research, 39,* 1321–1327.

Pomerantz, J. R., & Pristach, E. A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 635–649.

Raab, D. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences, 43,* 574–590.

Rajashekar, U., Cormack, L. K., & Bovik, A. C. (2002). Visual search: Structure from noise. In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications* (pp. 119–123). New York: ACM Press.

Ramachandran, V. S. (1988, August). Perceiving shape from shading. *Scientific American, 259*(2), 76–83.

Rao, R. P. N., Zelinsky, G. J., Hayhoe, M. M., & Ballard, D. H. (2002). Eye movements in iconic visual search. *Vision Research, 42,* 1447–1463.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9,* 347–356.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111,* 333–367.

Rensink, R. A., & Enns, J. T. (1995). Preemption effects in visual search: Evidence for low-level grouping. *Psychological Review, 102,* 101–130.

Rentschler, I., Hubner, M., & Caelli, T. (1988). On the discrimination of compound Gabor signals and textures. *Vision Research, 28,* 279–291.

Saarinen, J. (1996). Visual search for positional relationships between pattern elements. *Scandinavian Journal of Psychology, 37,* 294–301.

Sagi, D. (1995). The psychophysics of texture segmentation. In T. Papathomas, C. Chubb, E. Kowler, & A. Gorea (Eds.), *Early vision and beyond* (pp. 17–25). Cambridge, MA: MIT Press.

Smith, D. G., & Mewhort, D. J. K. (1998). The distribution of latencies constrains theories of decision time: A test of the random-walk model using numeric comparison. *Australian Journal of Psychology, 50,* 149–156.

Sternberg, S. (1966, August 5). High-speed scanning in human memory. *Science, 153,* 652–654.

Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction time experiments. *American Scientist, 57,* 421–457.

Sternberg, S. (1975). Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology, 27,* 1–32.

Stone, M. (1960). Models for choice reaction time. *Psychometrika, 25,* 251–260.

Sun, J., & Perona, P. (1996, January 11). Early computation of shape and reflectance in the visual system. *Nature, 379,* 165–168.

Takeuchi, T. (1997). Visual search of expansion and contraction. *Vision Research, 37,* 2083–2090.

Tanaka, K., Fukada, Y., & Saito, H. (1989). Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology, 62,* 642–656.

Tanaka, K., & Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology, 62,* 626–641.

Tavassoli, A., van der Linde, I., Cormack, L. K., & Bovik, A. C. (in press). An efficient technique for revealing visual search strategies with classification images. *Perception & Psychophysics.*

Thornton, T. (2002). *Attentional limitation and multiple-target visual search.* Unpublished doctoral dissertation, University of Texas at Austin.

Thornton, T., & Gilden, D. L. (2001). Attentional limitations in the sensing of motion direction. *Cognitive Psychology, 43,* 23–52.

Townsend, J. T. (1972). Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology, 25,* 168–199.

Townsend, J. T. (1974). Issues and models concerning the processing of a finite number of inputs. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 132–185). Hillsdale, NJ: Erlbaum.

Townsend, J. T. (1990). Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science, 1,* 46–54.

Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes.* New York: Cambridge University Press.

Townsend, J. T., & Wenger, M. J. (2004). The serial-parallel dilemma: A case study in a linkage of theory and method. *Psychonomic Bulletin & Review, 11,* 391–418.

Treisman, A. (1988). Features and objects: The 14th Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 40*(A), 201–237.

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,* 15–48.

van der Heijden, A. H. C. (1975). Some evidence for a limited capacity parallel self-terminating process in simple visual search tasks. *Acta Psychologica, 39,* 21–41.

van der Heijden, A. H. C., La Heij, W., & Boer, J. P. A. (1983). Parallel processing of redundant targets in simple visual search tasks. *Psychological Research, 45,* 235–254.

Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review, 7,* 208–256.

Verghese, P., & Nakayama, K. (1994). Stimulus discriminability in visual search. *Vision Research, 34,* 2453–2467.

Wang, D., Kristjansson, A., & Nakayama, K. (2005). Efficient visual search without top-down or bottom-up guidance. *Perception & Psychophysics, 67,* 239–253.

Ward, R., & McClelland, J. L. (1989). Conjunctive search for one and two identical targets. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 664–672.

Wickelgren, W. A. (1977). Speed–accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41,* 67–85.

Wolfe, J. M. (1992). "Effortless" texture segmentation and "parallel" visual search are not the same thing. *Vision Research, 32,* 757–763.

Wolfe, J. M. (1998a). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–73). Hove, England: Psychology Press.

Wolfe, J. M. (1998b). What can 1 million trials tell us about visual search? *Psychological Science, 9,* 33–39.

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences, 7,* 70–76.

Wolfe, J. M., & Bennett, S. C. (1996). Preattentive object files: Shapeless bundles of basic features. *Vision Research, 37,* 25–44.

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.

Wolfe, J. M., Chun, M. M., & Friedman-Hill, S. R. (1995). Making use of texton gradients: Visual search and perceptual grouping exploit the same parallel processes in different ways. In T. Papathomas, C. Chubb, E. Kowler, & A. Gorea (Eds.), *Early vision and beyond* (pp. 17–25). Cambridge, MA: MIT Press.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005, May 26). Cognitive psychology: Rare items often missed in visual searches. *Nature, 435,* 439–440.

Zelinsky, G. J., Rao, R. P. N., & Hayhoe, M. M. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science, 8,* 448–453.

Zenger, B., & Fahle, M. (1997). Missed targets are more frequent than false alarms: A model for error rates in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1783–1791.

(*Appendixes follow*)

Appendix A

Details of Implementation and Model Parameters

Parameter Ranges Used for Simulation

The parallel and serial models of multiple-target search (MTS) are parameterized by five independent variables each. To effectively assess the full predictive range of each model, we divided each of the five parameter ranges into a linearly spaced set of values. The set of parameters was defined over ranges that were broad enough to capture virtually any type of realistic data pattern that might emerge from experiment yet had a high enough resolution to afford reasonable fits to data. The particular ranges used in our simulations are given below. Note that although both the serial and the parallel models share the parameters *S*, *V*, and *D*, they do not occupy the same parts of the parameter space.

Parallel model: $S \in [0.12, 0.53]$, $V \in [1.00, 3.20]$,

$$D \in [1.00, 2.20],\ C \in [1.00, 4.00],\ \gamma \in [0.20, 0.90].$$

Serial model: $S \in [0.05, 0.20]$, $V \in [0.40, 1.50]$,

$$D \in [1.00, 1.50],\ \beta \in [0.00, 9.00],\ T_Z \in [150.00, 1{,}000.00].$$

The parallel model was resolved on the discrete $S \times V \times D \times C \times \gamma$ lattice with resolution $10 \times 10 \times 7 \times 10 \times 8$. The serial model was resolved on the discrete $S \times V \times D \times \beta \times T_Z$ lattice with resolution $10 \times 10 \times 5 \times 10 \times 10$.

For each quintuple of parameter values, we simulated 2,000 search trials for each of the nine MTS conditions so that all of the standard errors in response time (RT) and error rate were less than two time steps or one half of a percent, respectively. For example, to obtain estimates of the RT and error rate in a Set Size 2, single-target condition associated with a specific parallel model, we simulated 2,000 trials with two random walks (one based on target evidence, one based on distractor evidence) that drifted between response criteria. The drift, variance, asymmetry, and criterion levels were set by the specific quintuple being simulated. RT was estimated as the average number of steps until one of the walks was absorbed at the correct target-present boundary; error rates were estimated as the proportion of times both walks terminated on the relaxed distractor boundary. The full library of parallel and serial models was computed by rote estimates of RT and error rate for all nine MTS conditions across the range of each parameter. These libraries were used to compute goodness of fit to the data for each member of the test bed; there were 29 studies in all.

Resampling Data

To determine the robustness of models, we calculated error bars on all of the parameters and the chi-squares by resampling data. Each experiment consisted of nine conditions of target number and set size. Each condition generated two quantities, an observer-averaged RT and an observer-averaged error rate. Associated with the averages were standard deviations generated by observer variability. Each average and standard deviation sufficed to describe a sampling distribution of average data for a particular target number and set size—data that were consistent with those actually obtained. These sampling distributions were used to estimate parameter variability in our models.

The process of model parameter estimation began by drawing one sample from each of the 18 sampling distributions (9 RT and 9 error conditions) that defined the received average data in a given experiment. These 18 numbers were the results of a hypothetical experiment, a pseudoexperiment, consistent with the one actually run. Serial and parallel models were then fit to the pseudoexperiment, yielding estimates of capacity limitation, boundary movement, and so on. We also obtained an estimate of goodness of fit (chi-squares or equivalent log-likelihood in this context) from each pseudoexperiment. The procedure was repeated 100 times to permit reliable estimates of both the model parameters and their variability given the observed variability of the data. The ensemble of 100 pseudoexperiments also generated 100 matched chi-squares that could then be used to effect a simple paired *t* test to determine which model, serial or parallel, better fit the data. This procedure of model selection is more conservative than that obtained by computing a single chi-square on the observed data. We required that there be a significant difference in goodness of fit for model selection.

This manner of resampling does not take into account that speeds and accuracies are correlated in real reaction time experiments. Insofar as our parameter estimations are based on 18 independently varying quantities, the inferred variability is in fact much larger than would be realized in 100 actual experimental replications. This leads to a slight loss in resolution between serial and parallel models. The conclusions that we draw are therefore conservative and highlight the large differences that exist among different stimulus domains. That is, the data that serial models fit are in fact quite distinguishable from data that parallel models fit.

Rationality of Criterion Relaxation and the Parameter *C*

Here, we show that the parameter *C* that figures prominently in the parallel model of MTS effectively instantiates a manner of criterion relaxation that approximates rational decision making. Figure A1 compares the bounds of an ideal decision maker with the representative boundary structure used by the parallel model of MTS when the parameter *C* is set equal to two. The figure expresses evidence accumulation for the Set Size 2 case by plotting the joint state of the two random walks as an *x*,*y* coordinate in two dimensions. In this representation, the joint state of the walks begins at the origin (at Time 0) and evolves over time as a two-dimensional Brownian motion through the state-space. The black and gray lines represent the rational decision bounds for maintaining 95% accuracy (these bounds are derived in the section that follows). Whenever the joint state of the two random walks passes beyond the target-present bound, a rational decision maker will respond, "Target-present." Provided that the model and prior information are complete and that both are correctly specified, a decision maker using such a bound will be in error in responding "target-present" only 5% of the time. Similarly, whenever the joint state of the two evolving walks falls to the left or below the target-absent bound, a rational decision maker will be in error in responding "target-absent" only 5% of the time. The pair of gray dashed lines show the decision structure implemented by

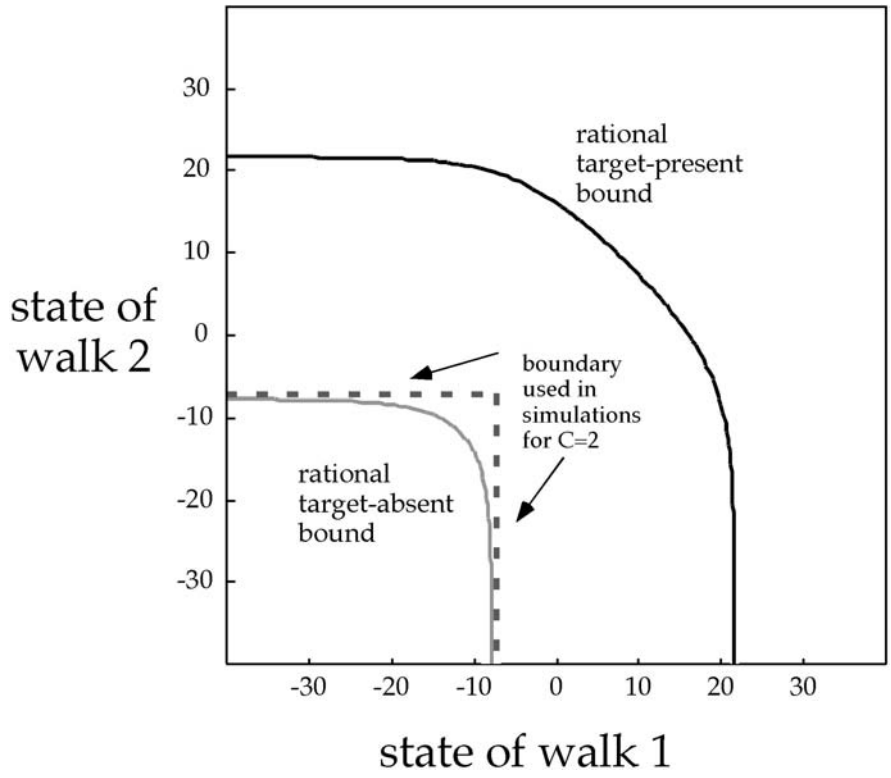

Figure A1. Similarity of the Set Size 2 decision bounds used by the parallel model of multiple-target search (MTS) and the optimal decision bounds of a rational observer. The abscissa and ordinate denote the current state of two random walks accumulating evidence on elements in a Set Size 2 display. The thick black curve denotes those points in the state-space where the posterior probability of the target-present hypothesis is .95. The lower gray solid curve denotes those points in the space where the posterior probability of the target-absent hypothesis is .95. The pair of gray dashed lines show the Set Size 2, target-absent criterion used by the parallel model of MTS when its criterion relaxation parameter C is set to two.

the parallel model of MTS for $C = 2$. The similarity of the bounds set by C and those of the optimal decision maker indicates that using a preponderance-of-the-evidence heuristic approximates a rational decision strategy.

The rational bounds are used to compute the patterns of RT and error expected of an ideal observer who makes decisions on the basis of the evidence accumulated in a set of random walks. These predictions are shown in Figure A2 for the case of parallel unlimited-capacity processing (no influence of set size on the random walks); predictions are shown for both a multiple-target (left panels) and a single-target (right panels) design. The ideal observer is constructed to maintain a constant average error rate of 5% across set size. Note that the RT and error predictions shown in the left panels of Figure A2
bear a striking resemblance to actual MTS data from Class A (see Figure 9 in main text). The target-absent RT predictions from the ideal observer decrease with set size at roughly the same rate as the pure-target conditions (the dashed line), indicating that the mirroring among these conditions is also a property of rational decision making. Recall that this kind of mirroring is common across the MTS data sets we have collected.

The errors of the rational observer are also similar to actual MTS data—the miss rates rise with set size, whereas the false alarms remain low. As the right panels of Figure A2 show, these structures emerge only in the context of a multiple-target design. When the rational observer is given a standard single-target design (i.e., no more than one target can appear at any set size), the predicted target-absent RTs will rise with set size at the same rate as the single-target RTs, and both the miss rates and false alarms will remain constant across set size at 5%.

Deriving Rational Decision Bounds in Visual Search

The rational bounds used in the previous two figures are based on standard expressions from statistical decision theory. These expressions provide the optimal rule for arriving at a decision given multiple lines of evidence (see Appendix IA in Green & Swets, 1988). The derivation begins with the posterior probability of the hypothesis that a display is of the target-present class (H_P) given the current evidence $\varepsilon(\tau)_1$ that has accumulated at time τ for Element 1:

$$P(H_P|\varepsilon(\tau)_1) = \frac{P(\varepsilon(\tau)_1|H_P)P(H_P)}{P(\varepsilon(\tau)_1|H_P)P(H_P) + P(\varepsilon(\tau)_1|H_A)P(H_A)} \quad \text{(A1)}$$

(*Appendixes continue*)

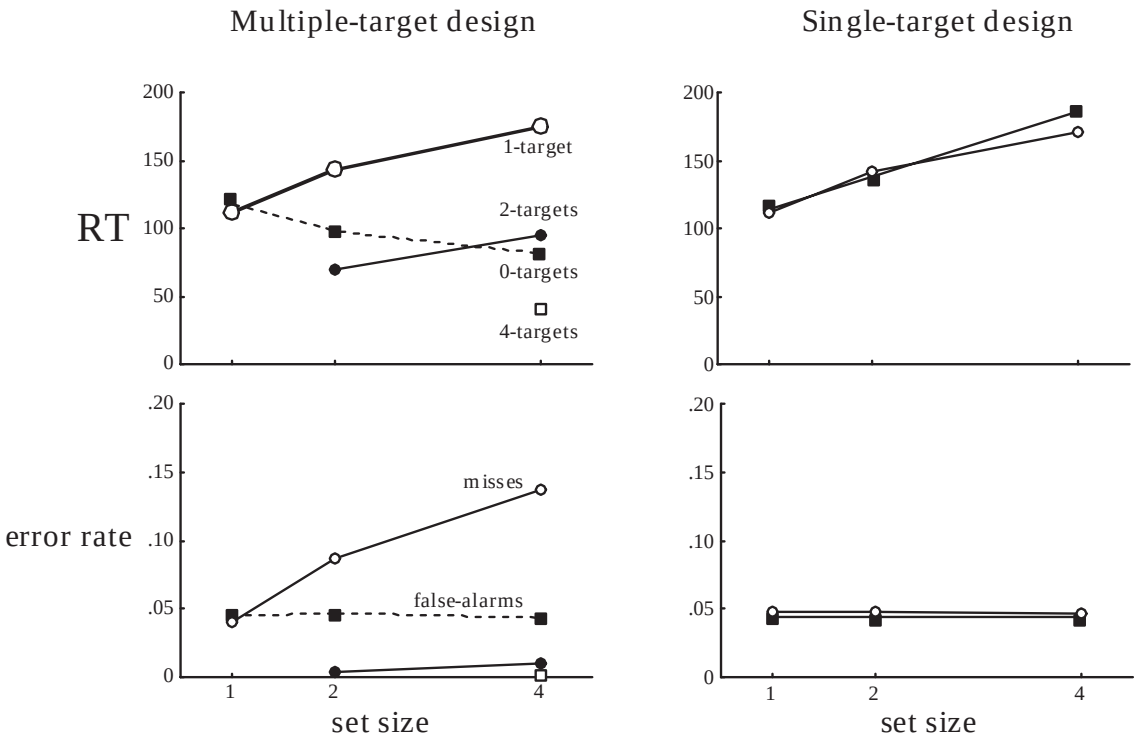

Figure A2. Predictions of an unlimited-capacity parallel model of search with rational decision bounds. The left panels show the predictions for response time (RT; upper left panel) and error (lower left panel) in a multiple-target search design. The right panels show companion predictions for the same model under a single-target design. Note that the legend in the upper left panel also pertains to the upper right panel, whereas the legend in the lower left panel also pertains to the lower right panel. Dashed line = target-absent trials.

(H_A denotes the alternative hypothesis of a target-absent display). By dividing out the numerator, we can reexpress Equation A1 as

$$P(H_P|\varepsilon(\tau)_1) = \frac{1}{1 + \dfrac{P(H_A)P(\varepsilon(\tau)_1|H_A)}{P(H_P)P(\varepsilon(\tau)_1|H_P)}}. \quad \text{(A2)}$$

The quantity $P(\varepsilon(\tau)_1|H_A)/P(\varepsilon(\tau)_1|H_P)$ is the likelihood ratio of the evidence given the competing target-present and target-absent hypotheses; the quantity $P(H_A)/P(H_P)$ is the ratio of the prior probability of the target-absent hypothesis to the target-present hypothesis.

We can extend Equation A2 to express the posterior probability of the target-present hypothesis given the current joint state of two walks:

$$P(H_P|\varepsilon(\tau)_1,\varepsilon(\tau)_2) = \frac{P(\varepsilon(\tau)_1,\varepsilon(\tau)_2|H_P)P(H_P)}{P(\varepsilon(\tau)_1,\varepsilon(\tau)_2|H_P)P(H_P) + P(\varepsilon(\tau)_1,\varepsilon(\tau)_2|H_A)P(H_A)},$$

where $\varepsilon(\tau)_1$ and $\varepsilon(\tau)_2$ represent the accumulated evidence at time τ for the random walks corresponding to Elements 1 and 2, respectively. Again, this expression can be simplified as

$$P(H_P|\varepsilon(\tau)_1,\varepsilon(\tau)_2) = \frac{1}{1 + \dfrac{P(H_A)P(\varepsilon(\tau)_1,\varepsilon(\tau)_2|H_A)}{P(H_P)P(\varepsilon(\tau)_1,\varepsilon(\tau)_2|H_P)}}. \quad \text{(A3)}$$

The quantity $P(\varepsilon(\tau)_1,\varepsilon(\tau)_2|H_A)/P(\varepsilon(\tau)_1,\varepsilon(\tau)_2|H_P)$ is the ratio of the likelihoods of the joint walk states given the competing hypotheses.

To compute the rational decision bounds for Set Size 2, Equation A3 must also incorporate aspects of the search design (i.e., both the prior probabilities of encountering various types of target-present and target-absent trials and the varied ways target and distractor elements can appear in such displays). For example, each conditional probability in the right-hand side of Equation A3 can be expanded to reflect a combination of the different trial types subsumed by each hypothesis. Consider the probability $P(\varepsilon(\tau)_1,\varepsilon(\tau)_2|H_P)$—this is the likelihood of the joint state of evidence given the hypothesis that the current trial is in the target-present class. For a Set Size 2, MTS display, there are three possible trial types: target–distractor (*td*), distractor–target (*dt*), and target–target (*tt*). Thus, we end up expressing the likelihood of $\varepsilon(\tau)_1$ and $\varepsilon(\tau)_2$ given H_P as

$$P(\varepsilon(\tau)_1,\varepsilon(\tau)_2|td)P(td) + P(\varepsilon(\tau)_1,\varepsilon(\tau)_2|dt)P(dt) + P(\varepsilon(\tau)_1,\varepsilon(\tau)_2|tt)P(tt),$$

where the prior probabilities of the display types—*P(td), P(dt),* and *P(tt)*—sum to H_P. By the assumption of independent evidence, this likelihood becomes

$$P(\varepsilon(\tau)_1|t)P(\varepsilon(\tau)_2|d)P(td) + P(\varepsilon(\tau)_1|d)P(\varepsilon(\tau)_2|t)P(dt) + P(\varepsilon(\tau)_1|t)P(\varepsilon(\tau)_2|t)P(tt). \quad \text{(A4)}$$

Expression A4, along with a similar expansion given the target-absent hypothesis, is plugged into Equation A3 and simplified using the actual prior probabilities of display types as provided by the experimental design.

In the end, we are able to express the posterior probability of the target-present hypothesis in terms of basic quantities—the likelihood ratio of each independent line of evidence and the probabilities with which different conditions occur in the design. The rational observer uses this exact expression to form decision bounds. The target-present bound is simply those points in the state-space (one-, two-, or four-dimensional depending on set size) where the posterior probability reaches 95% (the absent bound denotes those points where the probability reaches 5%). The final expression for the Set Size 1 case is the same for both single-target search (STS) and MTS designs and is given by

$$P(H_P|\varepsilon(\tau)_1) = \left(1 + \frac{B_1}{L_{\varepsilon(\tau)_1}}\right)^{-1}. \quad \text{(A5)}$$

The term B_1 in the numerator is the ratio of the prior probabilities of the competing hypotheses, $P(H_A)/P(H_P)$, and reflects any a priori bias in the model at Set Size 1 to favor one hypothesis over the other (for $B_1 > 1$, the model is biased to respond, "Target-absent"). The quantity $L_{\varepsilon(\tau)1}$ in the denominator is the likelihood ratio of the evidence in Walk 1 (cf. Equation A2). For a random walk model based on Gaussian increment distributions with symmetric means ($\pm S$) and common variance (V), the likelihood ratio reduces to $\exp(\varepsilon(\tau)_1 2S/V^2)$.

The exact expressions used for computing the Set Size 2 predictions in MTS (Equation A6) and STS (Equation A7) are provided below without further derivation.

$$P(H_P|\varepsilon(\tau)_1,\varepsilon(\tau)_2)_{MTS} = \left(1 + \frac{B_2}{\frac{1}{4}(L_{\varepsilon(\tau)_1} + L_{\varepsilon(\tau)_2}) + \frac{1}{2}(L_{\varepsilon(\tau)_1} L_{\varepsilon(\tau)_2})}\right)^{-1}, \text{ and} \quad \text{(A6)}$$

$$P(H_P|\varepsilon(\tau)_1,\varepsilon(\tau)_2)_{STS} = \left(1 + \frac{B_2}{\frac{1}{2}(L_{\varepsilon(\tau)_1} + L_{\varepsilon(\tau)_2})}\right)^{-1}. \quad \text{(A7)}$$

The term B_2 in the numerators is the a priori bias the model has at Set Size 2 to favor one hypothesis over the other. Comparison of Equations A6 and A7 shows that at Set Size 2, the MTS expression incorporates a product of the likelihood ratios in addition to a sum of the likelihood ratios.

The corresponding MTS expression for Set Size 4 is considerably more complicated as it entails combining the design probabilities with the multitude of ways targets and distractors can appear in the three distinct types of target-present conditions (i.e., *tttt* plus the permutations of *tddd* and *ttdd*). The Set Size 4 expression is given below without derivation.

$$P(H_P|\varepsilon(\tau)_1,\varepsilon(\tau)_2,\varepsilon(\tau)_3,\varepsilon(\tau)_4)_{MTS} = \left(1 + B_4\left(\frac{1}{12}\sum_{k=1}^{4} L_{\varepsilon(\tau)k} + \frac{1}{18}\left(\sum_{k=1}^{3}\left(L_{\varepsilon(\tau)k}\sum_{j=k+1}^{4} L_{\varepsilon(\tau)j,}\right)\right) + \frac{1}{3}\prod_{k=1}^{4} L_{\varepsilon(\tau)k}\right)^{-1}\right)^{-1}. \quad \text{(A8)}$$

The term B_4 is the a priori bias the model has at Set Size 4 to favor one hypothesis over the other.

The predicted patterns of RT and error shown in Figure A2 were generated by simulating a series of random walks appropriate to each trial type using the expressions appropriate to either an MTS or an STS design. For these simulations, all bias terms (B_k) were set to one (i.e., no bias). With each additional increment in time, the rational observer simply computes the posterior probability of the target-present hypothesis given the current joint walk state. The posterior probability is itself a random walk in probability space that begins at .5 (the ignorant state) and drifts over time between 0 and 1. The rational observer makes decisions by monitoring the status of the posterior probability: If this probability exceeds .95, the observer responds, "Target-present"; if this probability falls below .05, the observer responds, "Target-absent." The predicted RT by condition is the time required (number of steps) to reach one of these probability bounds. The associated error rate is the proportion of times the rational observer arrives at an incorrect conclusion (by design, this has to equal 5% averaged across conditions at each set size). Note that in MTS, error rates for some conditions can exceed 5% (e.g., misses for the Set Size 4, single-target cell), but these are offset by lower errors elsewhere so that the average target-present error remains invariant over set size.

Rationality of T_Z Relaxation

Much like the principled motivation behind the parameter β, the parameter T_Z in the serial model of MTS also instantiates a rational decision strategy. Whereas β is grounded in the increased prior probability of encountering distractor elements given large set-size displays (relative to the Set Size 1 case), T_Z approximates sensitivity to how priors get modified during the sequential analysis of a single *n*-element display.

The intuition behind the rationality of T_Z-based criterion relaxation arises in considering the following probability:

$$P(e_k \in D | e_j \in D, j < k \leq n). \quad \text{(A9)}$$

Here, e_k represents the currently uncategorized element under inspection, e_j is the previously categorized element(s), and D denotes the category of distractor. In words, this probability concerns the likelihood that the *k*th element is a distractor (prior to evidence accumulation) given that the previous element(s) was (were) categorized as distractors. If the previously identified distractor elements are somewhat certain to be in the distractor class D (i.e., categorization accuracy is high), then in the context of a typical visual search design, the prior for any unexamined element (e_k) will increase with the number of previous elements categorized as distractor. The intuition is that once the first element has

(*Appendixes continue*)

been categorized as a distractor, then it is somewhat more likely that the display in question is a target-absent display, and this of course implies that the remaining uncategorized elements are distractors. A rational observer would adjust his or her response criterion following each categorization so as to reflect the changing priors on the remaining elements. This kind of dynamic change to the response criteria is implemented in the model by T_Z relaxation.

In practice, we use T_Z to explore a range of possible relaxation rates in our simulations, recognizing that most of these values may not reflect optimal relaxation. One can approximate optimal boundary relaxation using analytic expressions that give categorization accuracy as a function of the number of accumulated evidence samples. These expressions depend on the basic model parameters (S and V), the sign and magnitude of accumulated evidence in the walk at time t, and the prior probability of each category. When the parameters are chosen such that categorization accuracy for single elements is close to perfect (i.e., the standard high-threshold regime; see Palmer et al., 2000), we are able to easily compute the probability in Equation A9 for the case in which the first element inspected in an n-element display has been categorized as a distractor. On the basis of the MTS design, the prior probability that the second to-be-inspected element will be a distractor, given that the first element has been categorized as such, is .80 for $n = 2$ and .86 for $n = 4$ (for a matched single-target design, these probabilities will be somewhat lower given the lack of redundant target displays).

Appendix B

Data and Model Fits

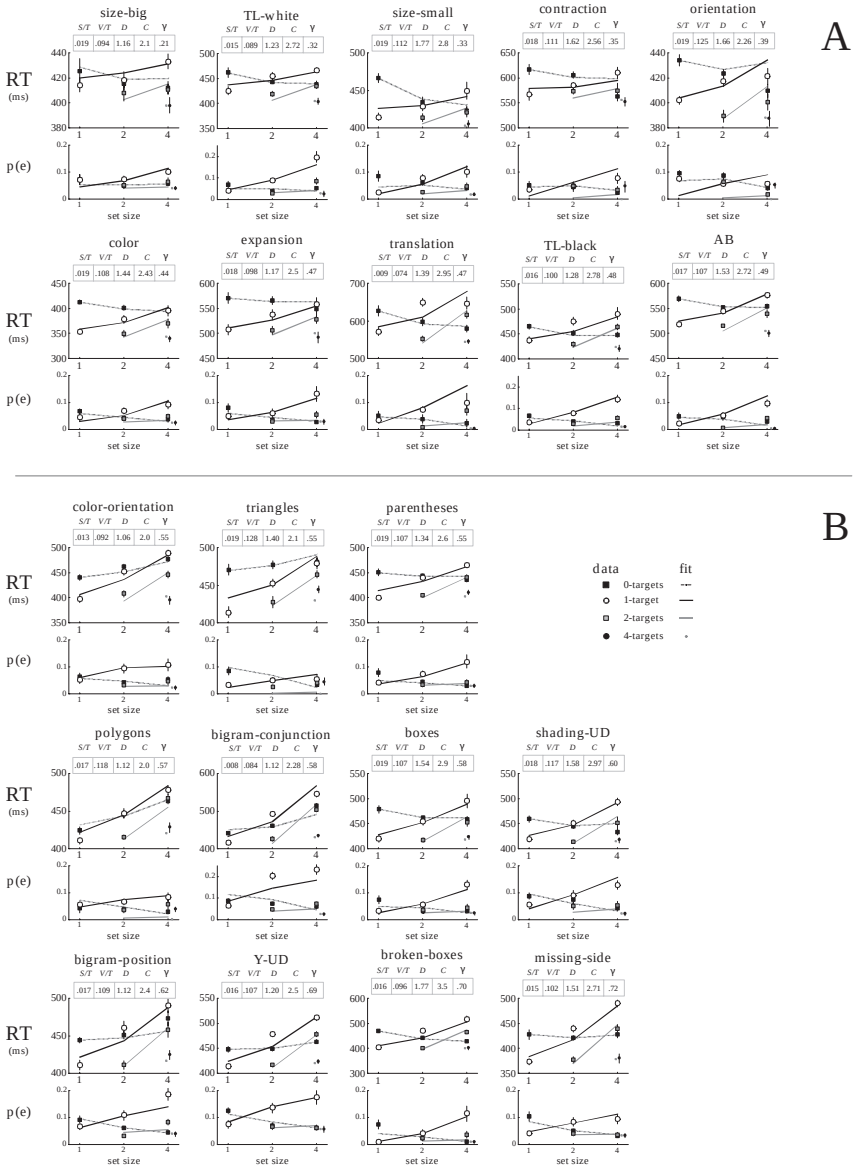

Figure B1. Data (points) and fits (lines) of the specific parallel model of multiple-target search that maximizes the likelihood of the data for each of the 21 tasks in Classes A and B. The 10 tasks in Class A are shown in the upper panel; the 11 tasks in Class B are shown in the lower panel. The parameter estimates based on fits to 100 resampled pseudoexperiments are inset for each task. RT = response time; S/T = average step size relative to target criterion; V/T = step-size variability relative to target criterion; D = criterion asymmetry; C = criterion relaxation parameter; γ = attention limitation parameter.

(*Appendixes continue*)

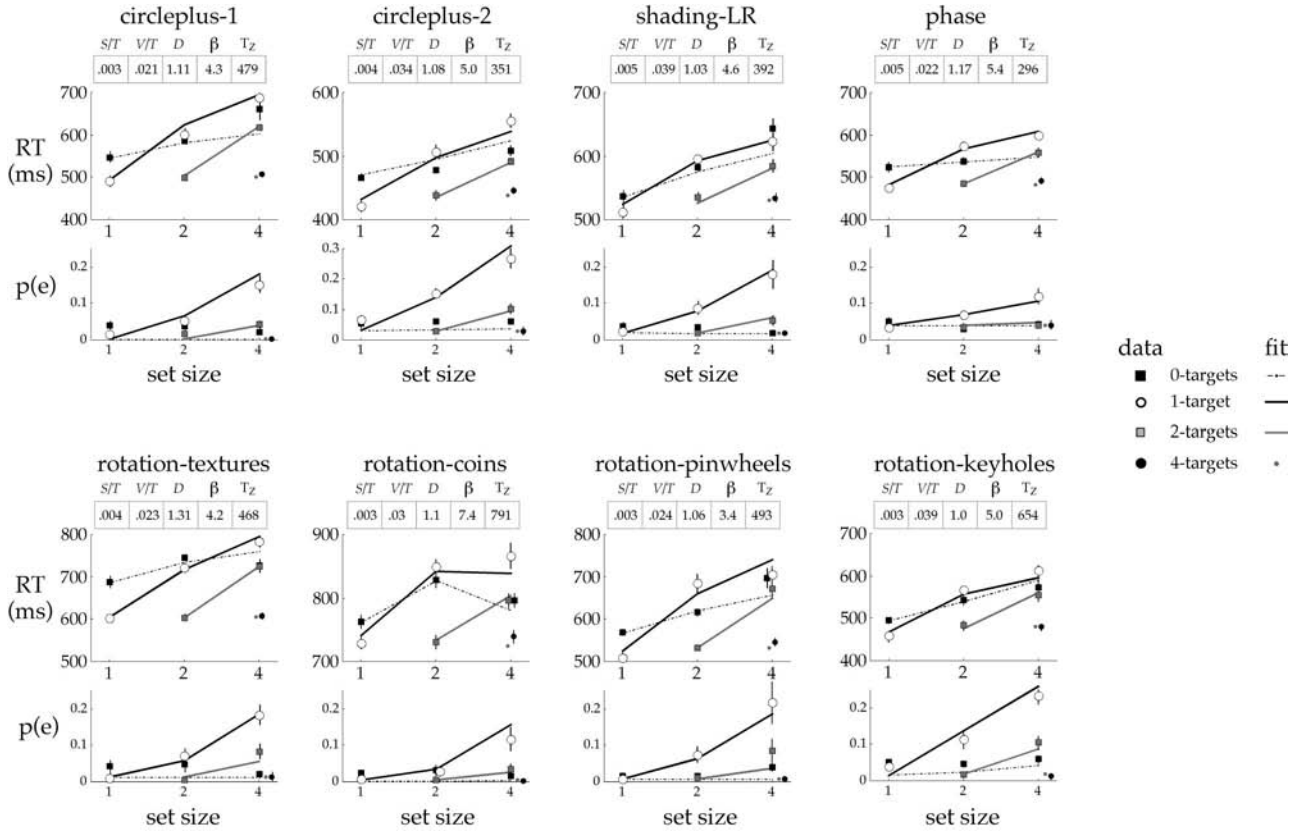

Figure B2. Data (points) and fits (lines) of the specific serial model of multiple-target search that maximizes the likelihood of the data for each of the eight tasks in Class C. The parameter estimates based on fits to resampled data are inset for each task. RT = response time; *S/T* = average step size relative to target criterion; *V/T* = step-size variability relative to target criterion; *D* = criterion asymmetry; β = set-size-dependent bias; T_Z = time-based relaxation parameter.

Appendix C

Parameter Tables

Table C1
Parameters Based on Best Fit of the Parallel Model to Class A and Class B Data Sets

	S/T		*V/T*		*C*		*D*		γ		χ^2-RT		χ^2-Err	
Task	Average	*SE*	Average	*SE*	Average	*SE*	Average	*SE*	Average	*SE*	Average	*SE*	Average	*SE*
Missing-side	.0151	±3e–4	.102	±.001	2.71	±.03	1.51	±.03	.72	±.010	2.87	±.10	1.42	±.07
Shading-UD	.0178	±2e–4	.117	±9e–4	2.97	±.03	1.58	±.02	.60	±.007	2.15	±.09	2.82	±.10
Color	.0187	±1e–4	.108	±7e–4	2.43	±.02	1.44	±.02	.44	±.008	1.59	±.07	2.17	±.10
Color-orientation	.0127	±3e–4	.092	±.001	2.02	±.01	1.06	±.01	.55	±.009	3.50	±.13	2.40	±.08
Boxes	.0191	±1e–4	.107	±7e–4	2.90	±.03	1.54	±.02	.58	±.007	1.64	±.09	2.04	±.09
Broken-boxes	.0159	±3e–4	.096	±.001	3.46	±.04	1.77	±.02	.70	±.008	3.51	±.12	2.23	±.08
Expansion	.0182	±2e–4	.098	±.001	2.49	±.04	1.17	±.02	.47	±.008	1.30	±.06	1.69	±.07
Contraction	.0176	±2e–4	.111	±.001	2.56	±.04	1.62	±.03	.35	±.010	3.44	±.14	4.96	±.12
A among Bs	.0166	±2e–4	.107	±9e–4	2.72	±.03	1.53	±.02	.49	±.007	2.46	±.10	4.18	±.13
Orientation	.0189	±1e–4	.125	±5e–4	2.26	±.02	1.66	±.01	.39	±.006	1.85	±.07	8.66	±.15
Parentheses	.0193	±1e–4	.107	±6e–4	2.62	±.03	1.34	±.02	.55	±.005	3.27	±.12	1.36	±.07
Size-big	.0192	±1e–4	.094	±6e–4	2.09	±.02	1.16	±.02	.21	±.003	2.01	±.13	2.15	±.08
Size-small	.0192	±1e–4	.112	±6e–4	2.79	±.03	1.77	±.02	.33	±.008	2.31	±.10	1.92	±.08
Translation	.0089	±2e–4	.074	±.001	2.95	±.04	1.39	±.02	.47	±.010	2.85	±.11	2.34	±.09
Triangles	.0190	±2e–4	.128	±9e–4	2.12	±.01	1.39	±.02	.55	±.006	3.97	±.15	5.51	±.12
Polygons	.0166	±2e–4	.118	±.001	2.03	±.01	1.12	±.01	.57	±.007	2.84	±.11	8.17	±.18
Bigram-conjunction	.0079	±1e–4	.084	±6e–4	2.28	±.02	1.12	±.01	.58	±.005	7.05	±.24	7.94	±.25
Bigram-position	.0167	±2e–4	.109	±.001	2.42	±.03	1.12	±.01	.63	±.006	1.99	±.09	3.28	±.12
Y-UD	.0159	±3e–4	.107	±.001	2.51	±.03	1.20	±.02	.69	±.009	4.29	±.16	1.42	±.05
TL-white	.0154	±3e–4	.089	±.001	2.72	±.05	1.23	±.02	.32	±.008	2.59	±.10	2.98	±.11
TL-black	.0157	±3e–4	.100	±.001	2.78	±.03	1.28	±.01	.48	±.008	1.89	±.09	2.31	±.08

Note. Averages and standard errors are based on fits to 100 pseudoexperiments. *S/T* = average step size relative to target criterion; *V/T* = step-size variability relative to target criterion; *C* = criterion relaxation parameter; *D* = criterion asymmetry; γ = attention limitation parameter; RT = response time; Err = error rate.

Table C2
Parameters Based on Best Fit of the Serial Model to Class C Data

	S/T		*V/T*		*D*		β/*DT*		T_Z		χ^2-RT		χ^2-Err	
Task	Average	*SE*	Average	*SE*	Average	*SE*	Average	*SE*	Average	*SE*	Average	*SE*	Average	*SE*
CirclePlus-1	.0031	±1e–4	.0209	±2e–4	1.11	±.009	.197	±.005	479	±12.0	2.00	±.09	7.24	±.18
CirclePlus-2	.0042	±1e–4	.0344	±6e–4	1.07	±.007	.236	±.005	351	±11.0	3.58	±.17	6.09	±.24
Shading-LR	.0052	±1e–4	.0389	±5e–4	1.03	±.005	.222	±.005	392	±11.8	1.77	±.08	1.84	±.07
Phase	.0053	±1e–4	.0216	±3e–4	1.17	±.010	.232	±.005	296	±8.8	1.04	±.06	1.68	±.07
Rotation-textures	.0041	±1e–4	.0230	±4e–4	1.31	±.013	.165	±.011	468	±17.0	1.69	±.09	4.61	±.11
Rotation-coins	.0033	±5e–5	.0296	±4e–4	1.10	±.006	.337	±.002	791	±16.0	1.68	±.07	2.99	±.11
Rotation-pinwheels	.0030	±1e–4	.0237	±5e–4	1.06	±.006	.161	±.006	493	±7.9	2.50	±.10	2.07	±.07
Rotation-keyholes	.0034	±1e–4	.0392	±8e–4	1.02	±.004	.249	±.004	654	±15.6	2.94	±.15	2.95	±.14

Note. Averages and standard errors are based on fits to 100 pseudoexperiments. *S/T* = average step size relative to target criterion; *V/T* = step-size variability relative to target criterion; *D* = criterion asymmetry; β/*DT* = set-size-dependent bias relative to distractor base criterion; T_Z = time-based relaxation parameter; RT = response time; Err = error rate.

Received April 29, 2005
Revision received September 6, 2006
Accepted September 7, 2006 ■