

RESEARCH ARTICLES

Eye movements in active visual search: A computable phenomenological model

T. D. KEECH AND L. RESCA

Catholic University of America, Washington, D.C.

We present a computational model and corresponding computer simulations that mimic phenomenologically the eye movement trajectories observed in a conjunctive visual search task. The element of randomness is captured in the model through a Monte Carlo selection of a particular eye movement based on its probability, which depends on three factors, adjusted to match to the observed saccade amplitude distribution, forward bias in consecutive saccades, and return rates. Memory is assumed to operate through tagging of objects already recognized as nontargets, which, in turn, requires their processing within the attentional area of conspicuity (AC). That AC is adjusted so that computer simulations optimally reproduce the distribution of the number of saccades, the failure rate for capturing the target, and the return rate to previously inspected locations. For their viability, computer simulations critically depend on memory's being long-ranged. In turn, the simulations confirm the formation of circulating or spiraling patterns in the observed eye trajectories. We also relate consistently the average number of saccades per trial to the saccade amplitude distribution by modeling analytically the combined roles of the AC in attention and memory. The full Supplemental Appendix A for this article may be downloaded from <http://app.psychonomic-journals.org/content/supplemental>.

Saccadic eye movements are made by humans and other higher animals in order to intermittently shift the high-acuity region of the eye retina, or fovea, to different areas of the visual scene. However, achieving a higher resolution is not the only reason why saccades are made. Saccadic eye movements may be influenced by cognitive demands—purposefully (endogenously) generated or stimulus (exogenously) driven, in any combination (Godijn & Theeuwes, 2002). In fact, it is also possible, under certain conditions, to inspect a scene adequately without making saccades, which is actually faster, although human subjects typically prefer not to do so (Zelinsky & Sheinberg, 1997). Studies of visual search have often been limited in the past—and still are, to some extent—to investigations of the *reaction times* (RTs) needed to find target stimuli. Further recording and studying of *eye movements* and *fixations* is supposed to provide more pertinent information and deeper understanding of underlying visual search and cognitive processes. Although those two types of measures are clearly related (Zelinsky & Sheinberg, 1997), one should always keep their distinction in mind when assessing corresponding studies and simulations.¹

In order to extract and characterize the basic elements of the visual processing of naturally complex scenes, scientists have also devised a host of visual search tasks in which fewer stimuli with more elementary, or *primitive*,

features are displayed to the viewing subject. A great deal about vision in general has indeed been learned from fundamental studies of such *paradigmatic* visual search tasks. On the other hand, one should always keep in mind the distinction and possible major differences that can occur between the results of studies and models of visual inspection in real-world scenes and settings (e.g., Canosa, 2005, 2009; Torralba, 2003; Torralba, Oliva, Castelano, & Henderson, 2006) and the results of studies and models of visual search through more or less artificial or remote representations or elements of those scenes (e.g., Najemnik & Geisler, 2005; Treisman & Gelade, 1980).

By exploiting major advances in information science and technology, computational models and simulations have been increasingly used in conjunction with empirical evidence in supporting or even formulating theoretical models of visual search in general and *active visual search* in particular.² In this article, we will present a contribution of ours to the latter. But first, we should mention at least some previous and most notable contributions to that literature and anticipate some of their relations or comparisons with ours.

That raises immediately the question of what is meant by *computational models* and how meaningful comparisons can be drawn between very different kinds of these. In a specialized psychoneural sense, a *bona fide computational*

L. Resca, resca@cua.edu

model may be one that tries to represent or mimic various computations presumably conducted by the brain while performing certain corresponding cognitive operations. A prototypical example of that is the computational model of visual attention of Itti and Koch (2001). To begin with, it computes the response of (lateral geniculate nucleus) neurons, tuned to center-surround intensity contrast, by convolving the luminance channel of the input image by a difference-of-Gaussians filter. It then proceeds similarly through higher and higher levels of visual processing, by computationally modeling the architecture presumed to inform, control, and deploy visuospatial attention. At the opposite extreme, a *computable phenomenological model* may attempt to merely reproduce through computer simulations the observed behavior of eye movements in a particular visual search task—for example, on the basis of a few empirical equations presumed to capture the essential functioning or manifestations of the major underlying psychoneural mechanisms. Our own model provides a primary example of that. Between or beside those modeling extremes, there are many kinds of variations or hybrids. For short, we may loosely refer to any of those as *computational models* in the generic sense of those words, but in each case, we will be careful to specify or classify more precisely what the model is or does.

Historically, it is appropriate to begin our brief survey with *guided search theory* (GST), which has been extensively simulated (Cave & Wolfe, 1990; Wolfe, 1994) for the following reason in particular:

Explicitness. Simulation of GS2 requires that underlying assumptions, parameters, and so forth be made explicit. A number of more qualitative models in the literature contain some intriguing ideas. There is nothing inherently wrong with qualitative models, but the models of visual search are becoming sufficiently complex so that the qualitative description may not capture their behavior. The parts of the models interact, sometimes in unexpected ways. Only by implementing the model can we be reasonably sure that it does what we say it does. (Wolfe, 1994, p. 233)

This describes just as well the intent of our effort, although a different kind of *source model* and eye-movement observations are simulated (Keech, 2006; Keech, Greenwood, Resca, & Parasuraman, 2009; Keech & Resca, 2010). As has already been mentioned, the difference is that our source model (Motter & Belky, 1998a, 1998b) is purely behavioral and phenomenological. Vice versa, GST at least starts with the computational realization of an *activation map*, as presumably generated by the brain, reflecting preattentive processes that combine bottom-up (similarities among items) and top-down (target specifications) information.³ That accounts for both kinds of feature and conjunction searches, set along a continuum. What differs across visual search tasks is the effectiveness of preattentive processes at guiding the attentional focus. Observations from a broad set of fairly disparate search tasks have been accounted for, rather than fitted, by GS2 simulations

on the basis of a fixed set of optimized parameters (Wolfe, 1994). On the other hand, GS2 simulations have been limited to RT predictions and relatively paradigmatic visual search tasks.

Greater emphasis on and development of a *saliency map* have appeared in the computational model of Itti and Koch (2000, 2001), which further mimics the underlying neural circuitry of visual attention (Itti, Koch, & Niebur, 1998; Koch & Ullman, 1985). That model is more closely inspired by consideration of the neural correlates of visuospatial attention, and it has been applied both to paradigmatic visual search tasks, such as those with feature pop-out and feature conjunction introduced by Treisman and Gelade (1980), and to more complex visual searches of natural scenes. Unlike our computer simulations, the Itti-Koch model attempts to predict search trajectories without any direct input from observations of actual eye movements. This type of modeling is evidently more fundamental and explanatory than ours. On the other hand, it was originally developed with some notable limitations. Mainly, bottom-up mechanisms of attention and inhibitory tagging were emphasized, but top-down mechanisms needed for object recognition or search planning were initially ignored. Distinctions between covert orienting of attention and eye movements were also blurred. Comparison between the performance of the model and that of actual human subjects was thus indirect.⁴

As an alternative to *item-based* activation maps, a smoother *area-based* activation map has been proposed and computationally implemented (Pomplun, Reingold, & Shen, 2003). Its purpose is to extend GST so as to better account for distractor grouping effects that figure more prominently in the *attentional engagement theory* (AET) of Duncan and Humphreys (1989, 1992). Conversely, the *area activation* model of saccadic selectivity of Pomplun et al. (2003) does not perform well in relatively sparse displays, where too many simulated saccades land between guiding items, rather than mostly on them, as observed empirically and assumed in our simulations, for example. Potentially more relevant to our own simulations are computational models of visual tagging that compare empirical with computer-generated scanpaths (Pomplun, Carbone, Koesling, Sichelschmidt, & Ritter, 2006).

The existence of scanpaths in eye movements, idiosyncratic to each individual subject viewing a complex scene, has long been established (Noton & Stark, 1971). This finding clearly challenges computational models whose aim is to predict human scanpaths based exclusively or primarily on bottom-up processes generated by a saliency or activation map (Itti & Koch, 2000, 2001; Pomplun et al., 2003). A saliency map, whether biologically based or not, can hardly account for the idiosyncratic biases of the viewing subject, which require further consideration of top-down processes, at a minimum. On the other hand, it is possible to compare *regions of interest* (ROIs) generated with bottom-up context-free image-processing algorithms (IPAs) with human ROIs, which further depend on top-down context-dependent representations of a scene (Privitera & Stark, 2000). That

can be done on a host of suitable measures, which thus indicate which IPAs treat an image more or less similarly to human viewers. Further considerations of top-down search templates for locating saliency peaks have been proposed (Rao, Zelinsky, Hayhoe, & Ballard, 2002; Tsotsos et al., 1995).

Stimulus *saliency*—hence, its modeling—must definitely play a basic role in object selection and guidance of eye movements, especially in free-viewing situations (Le Meur, Le Callet, Barba, & Thoreau, 2006; Parkhurst, Law, & Niebur, 2002; Peters, Iyer, Itti, & Koch, 2005). On the other hand, top-down influences, such as object *relevance* in relation to a more specific task, must also play a critical role—for example, in generating a *priority map* for object selection (Fecteau & Munoz, 2006). Various computational models have recently been extended to integrate bottom-up processing with top-down influences (Navalpakkam & Itti, 2002, 2006a, 2006b, 2007; Peters & Itti, 2007a, 2007b; Torralba, 2003; Torralba et al., 2006).

Integration of bottom-up processing with top-down constraints becomes even more imperative in natural environments in which subjects perform goal-driven *active tasks* coupled with viewing, such as making a cup of tea, washing hands, or walking down a hallway. Portable head-mounted eyetrackers have been developed to study eye movements outside of the laboratory, allowing demonstrations that saccade amplitudes and fixation locations and durations are highly dependent on the task that is being executed (Land, Mennie, & Rusted, 1999; Pelz & Canosa, 2001; Pelz, Canosa, & Babcock, 2000). Most notably (and understandably), fixations concentrate mainly on objects that are relevant to that task, rather than on visually salient but task-irrelevant stimuli. Canosa (2005, 2009) has correspondingly developed a biologically inspired computational model that combines bottom-up saliency elements with top-down task-related constraints into a *conspicuity map* that predicts more robustly the concentration of fixations in the task environment. This work fits into a broader framework and even more ambitious plan to model the brain as a computational system organized on a temporal scale and using *deictic* codes for the embodiment of cognition in performing complex tasks in real-world settings or simulations (Ballard, Hayhoe, Pook, & Rao, 1997).

As an alternative to heuristic computational models of visual search that can be applied to arbitrary stimuli (Itti & Koch, 2000; Rao et al., 2002), Najemnik and Geisler (2005) have developed “a formal, parameter-free analysis for a particular class of naturalistic stimuli” (p. 390), modeling the sensory and perceptual processing of an ideal Bayesian observer. For that ideal searcher, however, “the optimal eye movement strategy depends critically on how the visibility (detectability) of the target varies across the retina” (p. 387). That is, of course, the retina of real observers—that is, the human subjects in the experiments of Najemnik and Geisler. Now, the equivalent of the visibility maps or signal-to-noise ratio in Figure 2C in Najemnik and Geisler is precisely the *area of conspicuity* (AC) that Engel (1971, 1974, 1977) defined originally and many

other investigators have subsequently used, redefined, or further elaborated. See Motter and Simoni (2008) for a most recent investigation and a definitive review. The AC is a critical element also in our computational model. On the other hand, no other computational model has previously attempted to incorporate and account for the working of the AC in the way that either our model or that of Najemnik and Geisler has fundamentally done.

Najemnik and Geisler (2005) have also concluded that “simulations show that to achieve human performance levels it is necessary to have some coarse memory for past fixation locations so as to reduce the likelihood of returning to the same display region” (p. 390). They thus support *inhibition of return* (IOR), but only for one or two prior fixations. Our simulations indicate that some form of long-range memory is further involved in our type of visual search, which differs considerably from theirs. In Supplemental Appendix A, we discuss a possible origin for this apparent discrepancy.

Starting from the opposite idea of a *random walk* model of visual search (as we also did), Brockmann and Geisel (2000) have further assumed a power law dependence in the tail of the *saccade amplitude distribution* (SAD), for which they found empirical support in free viewing of natural scenes. Minimization of the time needed to scan the entire visual space then led them to predict that eye movement trajectories behave as *Lévy flights*, as opposed to more common *diffusive random walks*, which would result from a Gaussian SAD. But in order to obtain simulated eye trajectories that look like their observed scanpaths, an empirical determination of a *saliency field* for the correspondingly viewed scene is still needed. Brockmann and Geisel derived that saliency field from the spatial distribution of fixations made by observers throughout the scene (a picture of a party).

The stochastic assumptions of saccade generation made by Brockmann and Geisel (2000) involve *no memory*, consistent with an interpretation of visual search originally proposed by Horowitz and Wolfe (1998). However, the predictions and results of the Brockmann–Geisel model do not change substantially if those assumptions are relaxed so as to allow a sufficiently rapidly decaying correlation in the saccade sequences. Our simulations require a possibly coarse but long-range form of memory in order to be compatible with our observations based on a conjunctive visual search task, in contrast to their observations based on free viewing of natural scenes.

Various refinements of the basic Brockmann–Geisel model have been proposed by Boccignone and Ferraro (2004), particularly with regard to saliency field determination, measurement of scanpath complexity in terms of information entropy, and relations to scanpaths alternatively characterized (Privitera & Stark, 2000). Further evidence and characterization of Lévy-like diffusion in eye movements associated with spoken-language comprehension have recently been provided by Stephen, Mirman, Magnuson, and Dixon (2009).

Following another line of research, attempting to demonstrate that visual search has memory (or at least that

it can depend on some form of memory under relatively common or “normal” conditions), Peterson, Kramer, Wang, Irwin, and McCarley (2001) studied eye movements in a paradigmatic “T & Ls” conjunctive search task, quite similar or equivalent to those studied extensively by Motter and his colleagues (Motter & Belky, 1998a, 1998b; Motter & Holsapple, 2000, 2001, 2007; Motter & Simoni, 2007, 2008) and by ourselves (Keech, 2006; Keech et al., 2009; Keech & Resca, 2010). Essentially, they compared the observed proportions of item reinspections after various delays with those obtained by performing Monte Carlo simulations with three models: *memoryless*, *miss*, and *miss + realization*. Those comparisons allowed them to conclude that (with their stable displays) visual search has a memory of at least 12 items. The computational model that we present in this article is, to some extent, more fundamental and elaborate than theirs, since it takes explicitly into account other critical features of the visual search process, such as the SAD, the AC, and the *forward bias* in consecutive saccades. Nonetheless, with regard to the issue of long-range memory, their experiment, simulations, results, and conclusions are all quite similar to or compatible with ours.

A more basic model and computer simulations of visual search with variable memory (intermediate between sampling with and without replacement) were provided earlier by Arani, Karwan, and Drury (1984). As in most models but ours, they did not include any probability contribution for saccading to items as a function of distance from the current fixation, which is instead reflected in SADs. Also notable is an earlier modified random walk model and corresponding scanpath simulations used by Scinto, Pirlalamarri, and Karsh (1986) in a study of visual search in texture structures.

Following an alternative line of research, there have been many neural network (NN) or connectionist models of visual search. Some are interesting for particular reasons. For example, Cohen and Rupp (1999) have accounted for the transition from parallel to serial processing in visual search, and also for the role of *vertical symmetry*. Whereas they support revisions of *feature integration theory* (FIT; Treisman & Sato, 1990) and GST (Wolfe, Friedman-Hill, Stewart, & O’Connell, 1992) more than they support AET (Duncan & Humphreys, 1989, 1992), a SERR model of Humphreys and Müller (1993) does the opposite. A neural theory of attentive visual search that is an alternative to both FIT and GST has been advanced by Grossberg, Mingolla, and Ross (1994).

From our perspective, certain NN models are particularly interesting, since they may provide alternative views and underlying explanations of major features and findings of our phenomenological model. An adequate review and analysis of such models and issues are beyond the scope of this article. However, in a subsection of Supplemental Appendix A, we discuss, to some extent, the application of time series analysis and methods of complex dynamics in the study of active visual search (Aks, 2005, 2009; Aks, Zelinsky, & Sprott, 2002; Sprott, 2003) and related NN models (Kwok & Smith, 2005; Usher, Stemmler,

& Olami, 1995). At this point, we shall proceed with the presentation of our own computational model.

PURPOSE AND BASIC ELEMENTS OF THE MODEL

The purpose of our computable phenomenological model is to first identify from experiments the essential mechanisms that guide the conjunctive visual search and then express those essential mechanisms in a mathematical form and in computable terms so that we may understand and account for their interactions coherently and quantitatively, as opposed to merely describing those processes with words. In order to avoid the risk of getting immediately lost in its mathematical and numerical implementations, it is preferable first to discuss in general terms the working of the source model that we are trying to simulate numerically and its basic features that we are trying to represent mathematically.

Source Model

In our case, the visual search task is a paradigmatic one in FIT (Treisman & Gelade, 1980), where the target is characterized by a unique conjunction of color and orientation (black bar, tilted 45° to the right), among equal numbers of distractors that differ in either color (red) or orientation (tilted 45° to the left). The corresponding experiments were conducted originally with monkeys (Motter & Belky, 1998a, 1998b) and subsequently with human subjects (Keech, 2006; Keech et al., 2009). On the basis of their observations, Motter and Belky described the conjunctive visual search process in terms of a basic model, which mainly accounts for the role of the AC in guiding eye movements. Keech et al. expanded the scope of the Motter–Belky model by demonstrating the considerable presence of a mnemonic component.

The basic tenets and working of the Motter–Belky model, which we will try to simulate, are the following. (1) Recognition of the target typically occurs with high probability only within a restricted area surrounding the current fixation location. The size of this zone of attentional focus, which according to Engel (1971, 1974, 1977) is the AC, decreases with increasing density of relevant stimuli. (2) If the target is not detected within the current AC, a saccade is most likely made to an item just beyond that AC. That item most likely has the same color as the target, rather than the same orientation. (3) Except for those constraints, the item providing the new fixation point is selected apparently at random. (4) A forward bias is nevertheless apparent in consecutive saccades.

Some discussion of the literature on the AC has already been provided in the introduction to “Attentional Resources” in our companion paper (Keech & Resca, 2010). Here, more specifically for our purposes, we note that both *attentional* and *mnemonic mechanisms* are involved and deeply intertwined in how the AC works. Attentionally, the AC operates so that objects that currently fall within it can be effectively and simultaneously processed for target recognition, whereas objects that lie outside the current

AC cannot be sufficiently identified for that target recognition purpose.⁵ Now, as objects within the current AC are processed to a certain extent, the following can happen. If the target is recognized, the search essentially ends. If the target is not recognized, for whatever reason, two things happen. First, the inspected items are provisionally classified as *nontargets* or *distractors* and, correspondingly, retained in some sort of *memory store*. Second, an eye movement is made somewhat randomly to an item just outside or beyond the current AC. That item thus becomes the center of a new AC, comprising mostly, but not exclusively, not-yet-inspected items. Then the target recognition process starts anew. However, when the eyes are ready to jump (with a second saccade) to a third AC, there is also a tendency to avoid not only the preceding (second) AC, but also the (first) one that preceded that, in order to avoid, as much as possible, reinspections of items still accumulating in the memory store. This mechanism is called and understood by various authors⁶ as IOR, which thus has both an attentional component, related to previously attended ACs, and a mnemonic component, expressed as *inhibitory tagging* of items already inspected and still remaining in the memory store. As a result, IOR may also be responsible for a forward bias in consecutive saccades (Klein & MacInnes, 1999).⁷

At this point, we should immediately acknowledge that what generates or influences eye movements at a more fundamental or neural level is a question of much greater scope and complexity. Our computations will not attempt to address that question more precisely. Other, more “fundamental” computational models have tried to do so (e.g., Itti & Koch, 2000, 2001; Najemnik & Geisler, 2005). Rather, we attempt to (1) define mathematically and (2) implement computationally just the features and steps of the eye movements that we have described in the previous paragraph. Our purpose is thus to confirm, reject, or modify the Motter–Belky model, and to do so by determining whether simulated eye trajectories agree more or less with observed eye movements on various measures. Anticipating a major conclusion of ours, the memory store that is actually needed in order to reproduce various types of observations—and *return rates*, in particular—far exceeds, in terms of its capacity, current estimates associated with IOR in particular. So, the Motter–Belky model, augmented with some form of long-range memory, is what we shall call our *source model*.

Area of Conspicuity and Saccade Amplitude Distributions

The AC is alternatively known as the *functional visual field* (Motter & Simoni, 2008; Sanders, 1970), the *useful field of view* (Ball, Beard, Roenker, Miller, & Griggs, 1988), and the *visual lobe area* (Kraiss & Knäeuper, 1982; Widdel, 1983). It may be regarded as angularly symmetric around the foveation center (or point of fixation) within relatively uniform or random displays of stimuli, including regions near the display boundary. See the corresponding “sensitivity curves” in Figure 3B and the discussion of these in section 3.1.3 of Motter and Holsapple (2007). The

AC is then defined as a steeply decreasing probability of target recognition with increasing retinal eccentricity of the target (or its distance from the foveation center): See, for example, the “sensitivity curves” shown in Figure 5A in Motter and Belky (1998b) or, equivalently, the “signal-to-noise ratio” shown in Figure 2C in Najemnik and Geisler (2005). The AC also “zooms” as a function of stimulus spacing. Figure 5B in Motter and Belky (1998b) typically demonstrates this remarkable scaling property of the AC, which becomes essentially independent of array size once distances are expressed in average nearest-neighbor distance (ANND) units. Correspondingly, the AC decreases in size with increasing stimulus density, and the probabilities of target capture shown in Figure 5A in Motter and Belky (1998b) decrease overall with increasing number of stimuli in the display.

As captured (although not fundamentally explained) by Point 2 in the Motter–Belky model, there is an inherent relation between the AC, or its defining sensitivity curve, and the SAD observed phenomenologically. Namely, according to Point 2 of the Motter–Belky model, we expect that the maximum of the SAD corresponds roughly to the radius of the AC. But how could one define more precisely the AC radius? Perhaps one could fit the measured sensitivity curve with a Gaussian or a similar curve, determine its standard deviation, and then call that the AC radius. In our computational model, we take the following cruder approach, which is nonetheless adequate for our ultimate purposes. With regard to the *attentional* aspect of the AC, we will simply use the empirical SAD to operatively implement Point 2 of the Motter–Belky model. With regard to the *mnemonic* implication of the AC, which must be emphasized much more here than in the original Motter–Belky model, we assume an *equipotent* AC with a parameterizable radius. By that, we mean that if the target falls anywhere within the AC radius in a simulation, the target is recognized with certainty by our *computational observer*, and the simulation ends there. Conversely, as long as the target stays outside of the AC radius, the target is definitely considered as not-recognized, and the simulation proceeds to the next AC.

During the whole process, all the items that fall within each AC and are recognized as nontargets are added continuously to the memory store of our computational observer and are used to reduce the probabilities of fixating them in subsequent ACs. Now, starting with an initial estimate, the AC radius is varied as a parameter (together with other parameters of the model) until the computer simulations may eventually reproduce optimally the observed distributions of the number of saccades needed to capture the target, the observed failure rates to detect the target, and the observed return rates to previously inspected locations.

So, our equipotent definition of the AC amounts to approximating its correct probabilistic definition, represented by the observed sensitivity curve, with a step function that is 1 for target eccentricity within the AC radius (meaning certainty that the target is recognized therein) and 0 for target eccentricity beyond the AC radius (mean-

ing certainty that the target is not recognized therein). Of course, that is not exactly what is observed. For example, about 3% of the times that the target is even directly foveated, it is still not recognized (Motter & Belky, 1998a, p. 1808). Nevertheless, that is not enough to affect, overall, the main conclusions of our computational model, which, by the way, could be refined in principle by including the more precise probabilistic characterization of the AC through the observed sensitivity curves.

Forward Bias in Consecutive Saccades

A forward bias in consecutive saccades has been consistently observed and reported (Keech et al., 2009, Figures 10B and 11; Keech & Resca, 2010; Motter & Belky, 1998a, Figure 7B; Williams, 1966, p. 81). As has already been mentioned, it is likely a result of IOR, at least according to Klein and MacInnes (1999). An alternative view of IOR and forward bias has been advanced in terms of *attentional momentum*, attributing to attention some sort of “inertia,” reminiscent of that of linear momentum in Newtonian mechanics (Spalek & Hammad, 2004). In our investigation and computational model, we do not commit to either interpretation. Rather, we take into account the observed forward bias in consecutive saccades through an empirically determined probability factor of saccading consecutively in a forward direction.

Long-Range Memory

In visual search tasks with stable displays, various observations of low proportions of item reinspections, even after long delays, have consistently indicated the presence of some form of long-range memory (Dickinson & Zelinsky, 2007; Keech et al., 2009; Keech & Resca, 2010; Peterson et al., 2001). That memory persistence, observed even after as many as 12 lags in item refixations, considerably exceeds current estimates of IOR capacity, which appears to be limited to three to five spatially noncontiguous locations (Danziger, Kingstone, & Snyder, 1998; Snyder & Kingstone, 2000). Peterson et al. (2001) have suggested that their finding of an effectively larger visual memory capacity may be due to a grouped representation, or *chunking*, of items (see also Pashler, 1987). Dickinson and Zelinsky have more specifically, but not incompatibly, suggested a form of high-capacity but low-resolution *memory for the search path*. Had we known of such results and suggestions, it might have been possible, although not necessarily simple, to implement and test more directly similar ideas in our computational model. However, we originally designed and conducted our study independently and without any knowledge of either of those articles (Dickinson & Zelinsky, 2007; Peterson et al., 2001). So, we modeled a long-range memory contribution, which was indispensable for accounting for our observations as well, in a simpler (or simplest) form—that is, with a probability factor of IOR type, but with a long range, adjusted empirically so as to optimize our simulations of the observed eye movements.

Color Segmentation

The visual search task whose conduct we are trying to model and simulate phenomenologically is a paradigmatic

case in FIT (Treisman & Gelade, 1980) of presumably *serial conjunctive search*, where the target is characterized by a unique conjunction of two primitive features—that is, color and orientation. *Color segmentation* or *filtering* must figure prominently in that search (Egeth, Virzi, & Garbart, 1984; Pashler, 1987; Treisman & Gormican, 1988; Treisman & Sato, 1990). Indeed, distractors sharing target color are inspected a lot closer and at a much higher rate than are distractors sharing target orientation (Keech et al., 2009; Keech & Resca, 2010; Motter & Belky, 1998a, 1998b). In an extreme scenario, the latter, irrelevant distractors are rejected in parallel, whereas the former, relevant distractors are inspected sequentially. That extreme scenario is what we have actually assumed in our modeling. Of course, that is not exactly what happens. So, a fairly straightforward refinement of our modeling would have been to assume small, but nonzero, probabilities of fixating also near irrelevant distractors, while assuming a high probability, but not a certainty, of fixating only relevant distractors. These probabilities could have been extracted from corresponding observations—such as those shown in Figure 9 in Keech et al., for example. On the other hand, the overall effect of this refinement would have been fairly limited, having little or no consequence for the main conclusions (and purpose) of our computational model and simulations. So, for the time being, we ignored it.

In any case, it is useful to recall once more, in this context, that our model is not at all intended to predict, let alone explain, how attention is actually driven by environmental information, such as attentional capture by pop-out features, for example. So, we have hardly anything fundamental to say about either feature or conjunction searches in FIT or GST. All that we do is to model and simulate phenomenologically the behavioral data of a particular conjunctive search experiment. That sharply distinguishes our model from more general and explanatory models of neural activities and computations in visual search, such as those of Itti and Koch (2000, 2001) and others mentioned in the introduction, which do make predictions of attentional guidance through various saliency or activation maps.

Monte Carlo Procedure

In our computable model and simulations, we begin by assuming that only relevant objects (sharing target color) are precisely fixated and that the probability $p_{i \rightarrow j}$ of executing a saccade from a relevant object i to a relevant object j has the form

$$p_{i \rightarrow j} \propto f(r) g(r, \phi) h(l). \quad (1)$$

In Equation 1, the factors $f(r)$, $g(r, \phi)$, and $h(l)$ represent (not normalized) probabilities associated separately with the saccading distance, forward bias, and long-range memory involved in the incipient $i \rightarrow j$ saccade. The precise definitions and determination of these three critical factors will be discussed in the three following subsections. Here, we just mention that r represents the distance between the current fixation point on the i object and any other j object on which the incipient saccade may land,

ϕ represents the angle shift between the incipient saccade and the preceding saccade that landed on the i object, and l represents the number of intervening saccades from a possible previous inspection of the j object, starting with $l = 2$ for an immediate reinspection.

A computer simulation begins with the determination of a search array, which is generated as in the display seen by a subject in our experiment (Keech et al., 2009). Namely, the search items are positioned randomly over a $24.4^\circ \times 16.3^\circ$ field of view, with a minimum center-to-center separation of 1.4° . However, the target is always present in the virtual display, which was not always the case in the actual experiment. The initial (virtual) fixation is always located near the center of the display, at its closest relevant object, labeled as $i = 1$. From that—or subsequently, from any relevant object i corresponding to the current fixation—the distance d_{ij} to every other relevant object j in the display is computed. This distance d_{ij} is introduced in the distance factor $f(r)$ to obtain a corresponding probability. The angle shift ϕ_{ij} , ranging between the direction of the saccade just performed to the relevant object i and the direction of the potentially ensuing saccade to any other relevant object j , is then computed and associated to the j object—except when $i = 1$ corresponds to the initial fixation, of course. This angle shift ϕ_{ij} and the distance d_{ij} are thus introduced into the angle shift factor $g(r, \phi)$ to obtain a corresponding probability. Each relevant object within the current (equipotent) AC is then entered (with certainty) into the program memory of the computational observer, in order to be used in subsequent fixations to compute intervening $h(l)$ memory factors. Each relevant object is also evaluated to determine whether or not it has previously fallen within the AC of any preceding fixation. For relevant objects that have already been explored, the appropriate memory factor $h(l)$ is computed. For relevant objects that have not been previously explored, a factor of 1 is assumed in place of $h(l)$.

Now, the product of the distance, angle shift, and memory factors yields the relative probability of saccading to each relevant object from the current fixation point. Once computed for all relevant objects in the array, such relative probabilities must be normalized with a common factor K , so that their sum equals 1. The absolute probability of any possible $i \rightarrow j$ saccade then takes the final form

$$p_{i \rightarrow j} = K f(r) g(r, \phi) h(l). \tag{2}$$

Once all $p_{i \rightarrow j}$ s are so determined, corresponding segments are defined in the unit interval, and generation of a random number in that interval determines which j object is actually selected as the next fixation. This Monte Carlo procedure is then repeated from that point, until either the target is eventually captured by the program, by falling within an equipotent AC, or the computer search fails to find the target within the maximum allotted search time. Namely, failure is declared if the target, which is always present in our simulations, is not captured within 18 fixations (beyond the starting one). Such an upper limit of fixations is intended to match the experimental condition, in which a time limit of 4 sec was enforced on each trial. By dividing that time limit by the typical duration of a

fixation–saccade pair—namely, $4.0 / (0.200 + 0.030) \approx 18$ —we obtain our estimate of the maximum number (18) of fixations and saccades allowed on any simulated trial.⁸

Just to provide a simplified illustration, let us consider a hypothetical array comprising only four relevant objects, denoted as 1, 2, 3, and 4, with the target being Object 2. Suppose that the current fixation point is on Object 1 and that the probabilities $p_{1 \rightarrow j}$ have been evaluated as .4, .1, and .5 for $j = 2, 3,$ and 4, respectively. We may then assign to Object 2 a probability interval of $0 \leq \rho_{12} < .4$, to Object 3 a probability interval of $.4 \leq \rho_{13} < .5$, and to Object 4 a probability interval of $.5 \leq \rho_{14} \leq 1$. Now we generate a random number from a uniform probability distribution in the unit interval and compare it with those intervals. The interval within which the random number falls will select the saccade to the corresponding object. For example, suppose that the random number first drawn is .7812. Then, the selected saccade is to Object 4. At this point, the distances d_{4j} from Object 4 to all the other three objects are computed, along with the angles from Object 4 to Objects 1, 2, and 3. These angles are then subtracted from the angle from Object 1 to Object 4 to yield all three angle-shift ϕ_{4j} candidates. Assuming that only Object 1 fell within the AC of the previous fixation, a memory factor $h(2)$ is also assigned only to Object 1, with $l = 2$ corresponding to the two intervening saccades. The relative probabilities of saccading to each other object from Object 4 are thus computed, normalized, and used to assign probability intervals to each of those three objects. At this point, another random number is generated in the unit interval and is used to select the object of the ensuing saccade, according to which probability interval it selects. This process is repeated until either the target is captured or the upper limit of 18 iterations is reached. At that point, a new search array is randomly generated, and the next trial begins.

A final determination of the AC radius is reached as follows. At first, the AC radius is taken to coincide with the average distance between color-relevant nearest neighbors, or c-ANND. A set of simulations is then executed using this initial choice of AC. If two critical benchmarks of performance for the simulations—namely, the saccade number distribution and the failure rate for detecting the target—satisfactorily match the measured results, the data for that set of simulations are considered acceptable and recorded. Otherwise, the AC radius is modified by multiplying the c-ANND by an *extension coefficient* (typically, slightly greater than one, except for 12-distractor arrays: cf. Table 2). Another set of simulations is then generated and assessed, and so on, systematically modifying the extension coefficient until those critical benchmarks of performance for the simulations eventually satisfactorily match the corresponding measured data. The optimal AC radius and corresponding extension coefficient are thus obtained and recorded.

PROBABILITY FACTORS

In the previous section, we assumed that the probability of saccading from object to object in our phenomenologi-

cal modeling of a conjunctive visual search experiment depends on three elements—namely, the distance between search items, the angle shift between consecutive saccades, and some (additional) form of memory, presumably coarse but with a long range. Correspondingly, we introduced three basic factors in the construction of our computational model. These factors represent probabilities that either are extracted initially from experimental data, as in the case of $f(r)$, or reproduce finally experimental data through some parameter adjustments, as in the case of $g(r, \phi)$ and $h(l)$, which allow the computer simulations to reproduce optimally the observed eye movements. A critical element of randomness in the visual search is captured in the computer simulations through a Monte Carlo selection of a particular saccade, from one relevant object to another, based on the probability of that particular saccade. That probability, in turn, depends on the values assumed by the three factors for the considered consecutive fixations and their connecting and preceding saccades and fixations. Each of the angle shift $g(r, \phi)$ and memory $h(l)$ factors depends on two parameters, adjusted computationally to reproduce measured angle shifts and return rates, in particular. The memory factor $h(l)$ further depends on the size of the AC, which, in turn, is adjusted so that computer simulations optimally reproduce the observed distribution of the number of saccades needed to capture the target, the observed failure rate to detect the target, and the observed return rates to previously inspected locations. Complete color segmentation or filtering in parallel of the stimuli is assumed in the simulations, which is a simplification reasonably consistent with observations. In the next three subsections, we describe and discuss in greater detail the forms and features of these three factors that determine the simulated saccade probabilities.

Distance Dependence of the Saccade Probability

The distance factor $f(r)$ in the saccade probability $p_{i \rightarrow j}$ is obtained by dividing the measured SAD by the *interstimulus distance* (ISD) distribution. The average ISD distribution, shown in Figure 1, is obtained by sampling the distance from each search item to all other items over 10,000 randomly generated search arrays of all sizes. The ISD distribution is indeed independent of either array size or color segmentation. It begins at 1.4° , corresponding to the mini-

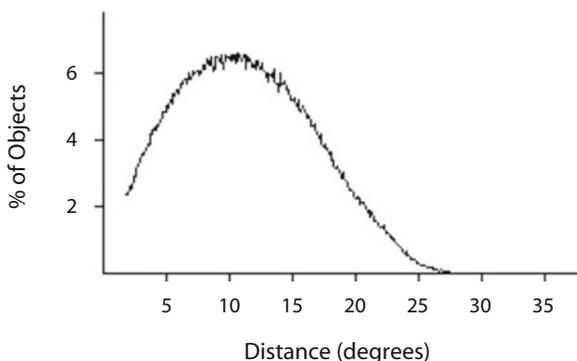


Figure 1. Average interstimulus distance distribution.

num center-to-center separation between search items, and terminates around 29° , corresponding to the diagonal of the rectangular display. In Appendix B we provide a more mathematical discussion of probabilities and distributions associated with $f(r)$, SAD, ISD, and their relations.

As has been noted elsewhere (Keech, 2006; Keech et al., 2009; Motter & Belky, 1998a), there is a qualitative difference in the saccade amplitude distributions for the first saccade and for midtrial saccades. In Figure 2, we show the saccade amplitude distributions for 48-distractor arrays for the first, second, and fourth saccades. One may notice some differences between young (panel A) and older (panel B) adults. In both groups, however, the amplitude distributions for the second and fourth saccades are relatively similar, and that is actually the case for any other midtrial saccade. On the other hand, the first-saccade amplitude distribution is more peaked and decays faster with distance, particularly for young adults. Therefore, at least two distinct probability distributions must be considered in our computational model—namely, one referring to the first saccade and another one representing all the subsequent midtrial saccades. We obtained the latter practically by averaging the measured second, third, and fourth SADs. We may refer to Aks (2009) for a broader perspective on the functional dependence of SADs and for further reference to studies of characteristic differences between initial and subsequent saccades.

The probability factors $f(r)$ used in the computational model (obtained by dividing the SADs by the ISD distribution) for the first saccade and the average of the second, third, and fourth saccades are shown in Figure 3 for young (panel A) and older (panel B) adults searching 48-distractor arrays. One may note a somewhat greater probability that young adults saccade farther away than older adults, which is consistent with young adults' having slightly greater ACs, as has been suggested elsewhere (Keech, 2006; Keech et al., 2009).

Angle-Shift Dependence of the Saccade Probability

For saccades with amplitudes shorter than about 7° , which corresponds roughly to the average saccade amplitude, there is a bias to saccade in a forward direction, relative to that of the preceding saccade. This is illustrated in Figure 11 in Keech et al. (2009), which shows that the highest concentration of shorter saccades occurs for relatively small angle shifts on either side of the direction of the preceding saccade. On the other hand, for longer saccades, a more balanced concentration occurs also around the backward direction.⁹ We may account phenomenologically for this angle shift dependence of the saccade probability $p_{i \rightarrow j}$ with a partly Gaussian and a partly constant factor—namely,

$$g(r, \phi) = \begin{cases} 1 + G_S \exp\left[-(\phi / \phi_S)^2\right] & \text{for } r < 7^\circ \\ 1 & \text{for } r \geq 7^\circ. \end{cases} \quad (3)$$

In Equation 3, ϕ denotes the angle shift between consecutive saccades, expressed in radians between $-\pi$ and π . After completion of our computational model and perfor-

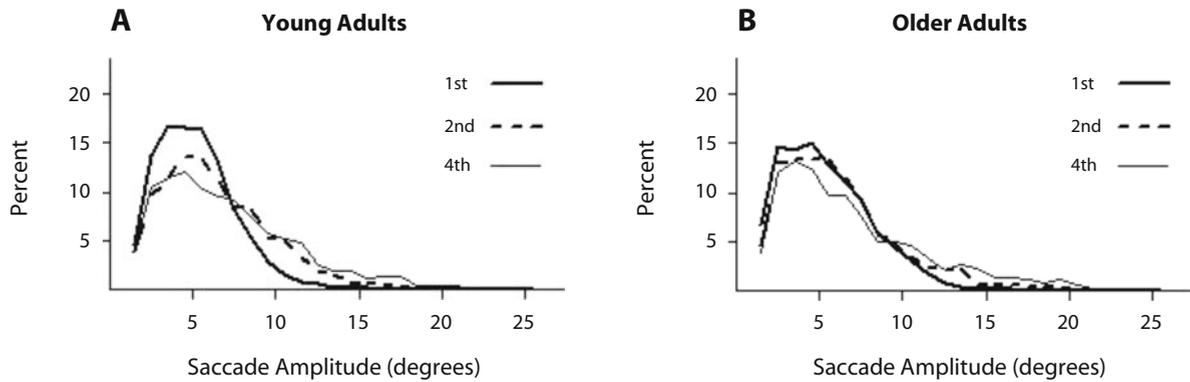


Figure 2. Saccade amplitude distributions for the first, second, and fourth saccades for 48-distractor arrays, for young (A) and older (B) adults.

mance of extensive computer simulations of the eye trajectories, we determined that the G_S parameter can be optimally adjusted to the values $G_S = 0.5, 1, 5,$ and 5 for both young and older adults searching arrays with 12, 24, 48, and 96 distractors, respectively. This trend of increasing G_S values reflects, at least in part, a relatively greater manifestation of the forward bias in larger array sizes, which is allowed by a correspondingly greater delay in reaching the display boundary from the initial central fixation. This, in turn, is consistent with observations of saccade direction cosine autocorrelations, as in Figures 1A–1D in Keech and Resca (2010). On the other hand, the ϕ_S parameter can be adjusted optimally to one radian (57.296°) in all cases.¹⁰

Memory Dependence of the Saccade Probability

Normalized return rates indicate fairly low probabilities of return to locations previously inspected, no matter how early (see Figure 20 in Keech et al., 2009). This suggests some form of long-range memory, which we tried to represent with (inverse) power laws in a corresponding memory factor in the saccade probability $p_{i \rightarrow j}$ of our computational model—namely,

$$h(l) = \frac{R_l}{(l - 1)^n} \tag{4}$$

In Equation 4, l represents the number of intervening saccades, starting with $l = 2$ for an immediate return, and the coefficient R_l and the exponent n are treated as two adjustable parameters. After completion of our computational model and performance of extensive computer simulations of the eye trajectories, we determined that the observed return rates and other key measures are optimally reproduced by selecting the exponent $n = 3$ for 12- and 24-distractor arrays and the exponent $n = 1$ for 48- and 96-distractor arrays, for both young and older adults. On the other hand, optimal values of the R_l coefficient are listed in Table 1.

The empirically determined parameter values imply the following. Since the probability of reinspection associated with the $h(l)$ factor decreases with longer delays, or larger l , that contributes more to primacy than to recency effects.¹¹ Thus, memory has a considerable long-range contribution (like a primacy effect). Also, on account of the values of both the R_l and n parameters, there is a lower probability of reinspection, especially after longer delays (larger l), in the smaller than in the larger arrays. This agrees with the smaller normalized return rates that are observed, overall, in the smaller arrays (see Figure 20 in Keech et al., 2009). Thus, there is more “memory” in searching smaller array sizes, although that is at least partly a consequence of the eye trajectories’ having, on average, fewer saccades in smaller arrays.

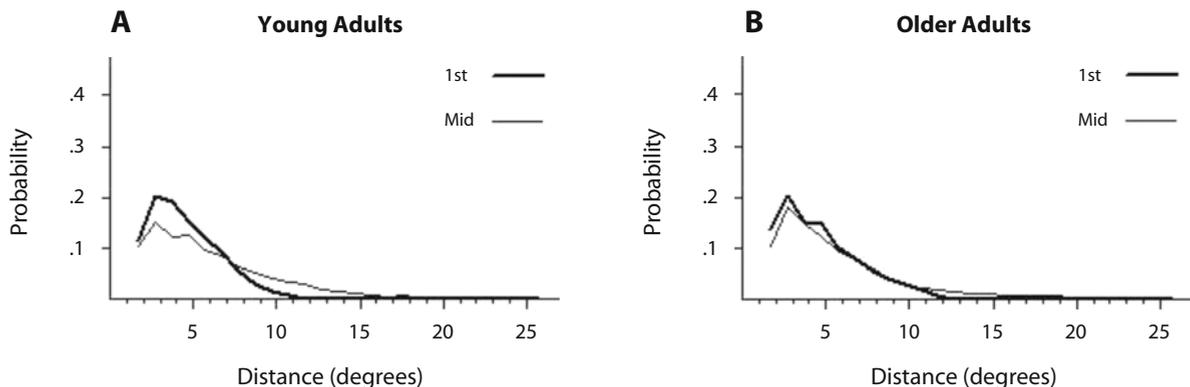


Figure 3. Saccade probability as a function of distance, or the $f(r)$ factor, for the first saccade and midtrial saccades for 48-distractor arrays, for young (A) and older (B) adults.

Table 1
Optimal Values of the R_I Coefficient for Young and Older Adults Searching Arrays of Each Size

Search Array Size	R_I Coefficient	
	Young Adults	Older Adults
12	0.004	0.005
24	0.008	0.008
48	0.060	0.050
96	0.125	0.100

In any case, by noting that the R_I values reported in Table 1 are much smaller than 1, whereas $h(l) = 1$ is assumed for items that have not yet been inspected, we must conclude that memory influences the search very strongly—a conclusion that will be later confirmed and further demonstrated by comparisons with *memoryless* simulations.

We should remark at this point that, although the memory factor $h(l)$ in the saccade probability $p_{i \rightarrow j}$ is closely related to normalized return rates in eye movement observations, one should not require or expect a precise correspondence, let alone a numerical fit, between the two. Indeed normalized return rates, like any other experimental observation, do not reflect exclusively the memory component, but also its interactions with other psychoneural mechanisms and geometric-spatial constraints that are involved in this particular visual search task (Keech & Resca, 2010). In our computational model, we attempt to isolate and characterize those mechanisms and constraints by attributing to each one a mathematical expression or a computational procedure. However, direct comparisons can be made only between observations and *final* products of the simulations, rather than between observations and individual components that are initially devised in the model construction. Ultimately, it is only a final product agreement that may give us confidence that individual components of the model capture at least some essential elements of the mechanisms that underlie the observations, thus permitting some further understanding and interpretation of those.

We should also remark that an optimal fit of simulated normalized returns to observed normalized returns equivalently implies an optimal fit of simulated plain returns to observed plain returns and conversely, because the number of opportunities to return depends only on the number of intervening saccades, whether simulated or observed (Keech et al., 2009).

Finally, we should report that at the beginning of our investigation, we tried to represent $h(l)$ with *direct* power laws, proportional to $(l - 1)^m$ with some positive m integers, for l up to a maximum of 5 or 6. That was assumed to represent only short-range memory or “conventional” IOR. However, we were not successful in generating acceptable simulations on that basis alone. So, we had to switch to inverse power laws to represent $h(l)$, which implies long-range memory.

COMPUTER SIMULATIONS

The viability of our computational model depends on a reasonable agreement between simulated and observed eye

movement trajectories on major statistical characteristics, which include (1) distributions of the number of saccades needed to capture the target, (2) failure rates for detecting the target, (3) saccade amplitude distributions, (4) saccade direction cosine autocorrelations, and (5) distributions and scatterplots of measures of circulation, C , and subtended area, A . In the following two subsections, the results of simulations are presented, and their abilities to match observations on those characteristics are assessed.

Simulation Results

Table 2 summarizes measured and simulated results, including averages and standard deviations, for some key performance measures.

Simulations were repeated in sets of 630 trials for each array size, in order to match the measurements for 90 target-present trials by 7 subjects, in each age group of the corresponding experiment (Keech et al., 2009). Since the inverse square root of 630 is about 4%, and standard deviations are found to be about half the corresponding means for both the saccade number and amplitude distributions, one expects a relative standard error of about 2% for those means. Indeed, repeating sets (i.e., samples) of 630 simulations with the model parameters that produce the closest matches to the measured data, we found that the average number of saccades needed to capture the target remains within 2% of the means listed in Table 2. Other measures of performance of the simulations involve the same order of relative stability. By the same token, an agreement within 2% or so between measured and simulated averages indicates a satisfactory match. So, we note in Table 2 that this is the case for saccade average numbers for all array sizes, but not so for saccade average amplitudes for any array size. We shall return to that shortcoming of the model at the end of this subsection and in the Discussion section.

It is also worth noting that, in Table 2, the extension coefficient is increasingly greater than 1.0 for 24-, 48-, and 96-distractor arrays, implying that multiple objects are entered in memory during each fixation for those array sizes. On the other hand, for 12-distractor arrays, the average number of saccades required to capture the target can be matched to measured data only by assuming an extension coefficient smaller than 1.0. That implies that only the fixated object can be inspected for target discrimination, and memorized if it is not the target, as the search continues. From an analysis of their data, Motter and Belky (1998a) also suggested the possibility that “the small number of potential targets in the small arrays may shift the efficiency of the search toward an item-by-item basis” (p. 1810).

Figure 4 shows the simulated saccade number (panel A) and saccade amplitude (panel B) distributions for 48-distractor arrays searched by older adults. The simulated saccade direction cosine autocorrelation is shown in Figure 5, and Figure 6 shows the simulated circulation distribution (panel A) and area–circulation scatterplot (panel B). With some exception, again, for the saccade amplitude distribution in Figure 4B, the simulation results

Table 2
Summary of Simulated (S) and Measured (M) Results

Array Size	Radius of Area of Conspicuity, in Degrees (Extension Coefficient)		Average Saccade Number (SD)		Average Saccade Amplitude, in Degrees (SD)		Search Failure Rate (%)	
	Young	Older	Young	Older	Young	Older	Young	Older
12 S	4.83 (<1)	4.75 (<1)	3.4 (1.9)	3.4 (1.9)	8.0 (4.2)	8.2 (4.2)	0	0
12 M	–	–	3.4 (2.0)	3.3 (2.0)	6.3 (3.4)	6.2 (3.5)	0	0
24 S	4.05 (1.1)	3.78 (1.004)	4.0 (2.2)	4.0 (2.3)	8.1 (4.1)	8.1 (4.1)	0	0
24 M	–	–	4.0 (2.3)	4.2 (2.4)	6.5 (3.4)	6.4 (3.7)	0	0
48 S	3.56 (1.32)	3.25 (1.20)	6.1 (3.5)	6.5 (3.8)	7.5 (3.9)	7.6 (4.2)	0	0
48 M	–	–	6.1 (3.7)	6.5 (3.8)	6.7 (3.6)	6.6 (3.9)	0.9	2.1
96 S	3.48 (1.64)	3.13 (1.43)	7.8 (4.6)	8.5 (5.0)	7.1 (3.7)	6.9 (3.9)	3.5	11.4
96 M	–	–	7.8 (4.7)	8.4 (4.7)	6.4 (3.4)	6.5 (4.0)	7.8	15.1

agree qualitatively with those of corresponding observations, which are superimposed in Figures 4A and 4B and referenced in the captions of Figures 4–6.¹²

Simulation results corresponding to those shown in Figures 4–6 for all array sizes searched by both young and older subjects have been generated and compared with the corresponding observations, providing the same quality of agreement (Keech, 2006). That already leads to the following conclusions. First, a qualitatively good agreement can be consistently attained with regard to saccade number distributions for both young and older adults and for all array sizes. Second, the distributions of the circulation, *C*, and its scatterplots with the area measure, *A*, computed for the simulated trajectories, follow basic trends that are the same as those observed for the cor-

responding eye movements. These results derive substantially from the following two basic factors that drive the simulated trajectories and mainly reflect two corresponding psychoneural mechanisms that underlie the observed eye movements. The first one is *h(l)*, reflecting long-range memory, and the other is the angle shift factor *g(r,φ)*, which reflects an additional forward bias in consecutive saccades (presumably related to short-range memory or “conventional” IOR). These two factors critically contribute to a tendency toward saccading around the periphery of the display, once that boundary has been reached by the eye trajectory within an AC radius or so, rather than returning to previously inspected central regions. Thus, *spiraling* or *circulating patterns* in the eye trajectories appear prevalently in both simulations and observations (Keech

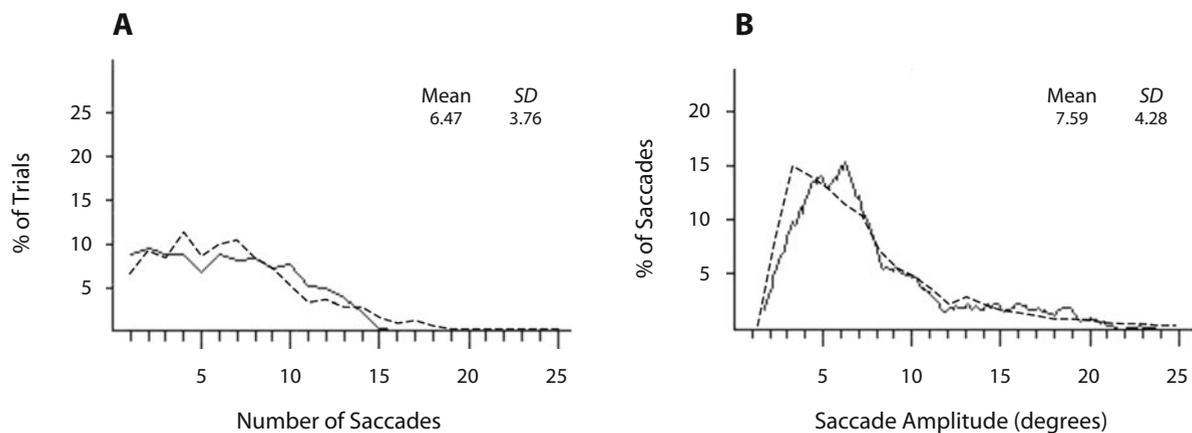


Figure 4. Simulated saccade number (A) and saccade amplitude (B) distributions for 48-distractor arrays searched by the older adults. The means/standard deviations are (A) 6.47/3.76 and (B) 7.59/4.28. The corresponding observed distributions are shown as thin lines in Figures 2C and 7C in Keech, Greenwood, Resca, and Parasuraman (2009) and are superimposed as broken lines here. Their means/standard deviations were (2C in 4A) 6.5/3.8 and (7C in 4B) 6.64/3.92.

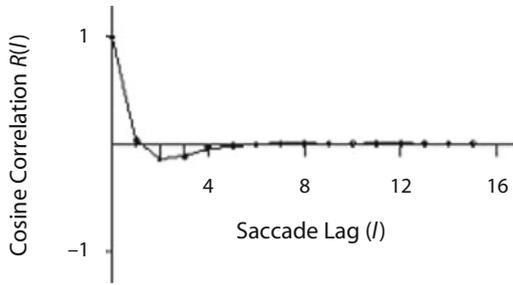


Figure 5. Simulated saccade direction cosine autocorrelation $R(l)$ corresponding to 48-distractor arrays searched by older adults. The experimentally observed autocorrelation is shown in Figure 1C in Keech and Resca (2010) on an expanded scale. A standard error of about 0.0012 was estimated for the latter in Appendix A in Keech and Resca (2010). This also implies that “error bars” are completely negligible on the scale of this Figure 5 or Figure 7 later, and differences in $R(l)$ values as small as .0012 should be regarded as meaningful.

& Resca, 2010). This interpretation is further supported by the study of saccade direction cosine autocorrelations in simulations. For example, $R(l = 1)$ is slightly positive in Figure 5, indicating forward progression in consecutive saccades, whereas $R(l = 2, 3, 4)$ values are negative, indicating 90° – 270° turning of direction in subsequent saccades, caused by reaching the display boundary. The same results are observed experimentally (cf. Figure 1C in Keech & Resca, 2010).

On the other hand, equally good agreement cannot be attained between simulations and experimental observations with regard to SADs, especially for the smaller array sizes. In particular, measured SADs exhibit a greater number of saccades with amplitudes shorter than the AC radius than is predicted by the computational model. This originates, at least in part, from the modeling assumption that all relevant objects within an AC must be entered in memory, which greatly reduces the chance that subsequent saccades may target objects already fallen within previous ACs, particularly for the smaller array sizes (see also the previous discussion of the values of n and R_l parameters

in Equation 4 and Table 1). Deactivating memory, which corresponds to setting $h(l) = 1$ in the simulations, reestablishes the match with the observed SADs. On the other hand, elimination of memory is unacceptable, as will be demonstrated more specifically later by comparisons with memoryless simulations. This current limitation of the model in consistently accounting for the interaction between attention and memory through the AC modeling will be discussed further in the Discussion section.

Sensitivity to Parameter Selections

In order to investigate the sensitivity of key measures of simulation performance to parameter values, a series of 630-simulation sets were executed with the AC extension coefficient, the angle shift factor, and the R_l coefficient alternatively varied. In each instance, only one parameter was varied, and the others were kept at the values found for matching the measured data optimally. These additional simulations were conducted using 48-distractor arrays and the model calibrated for the older adults, using their corresponding distance factor, $f(r)$. For each variation of one parameter, three independent 630-simulation sets were executed, thus forming a threefold repetition for each given parameter change. So, the overall results show the main effects of parameter variations on the model predictions, whereas the threefold repetitions confirm, in particular, the stability of key performance measures with respect to sampling distributions for 630-simulation sets (i.e., samples). The results are shown in Table 3, except for the circulation, C , which is found to remain relatively stable under all the considered variations.

The first five main rows in Table 3 show the variation in key performance measures as the AC extension coefficient (which generates the AC radius through multiplication by c -ANND) is varied from 1.0 to 1.4 in increments of 0.1. In correspondence to Figures 4–6 in particular, the third (shaded) main row shows the results of three independent 630-simulation sets where all the model parameters have values that optimize matching to all the measured data. This Set Group 3 thus corresponds to the model’s best performance, or *baseline*. As the AC extension coeffi-

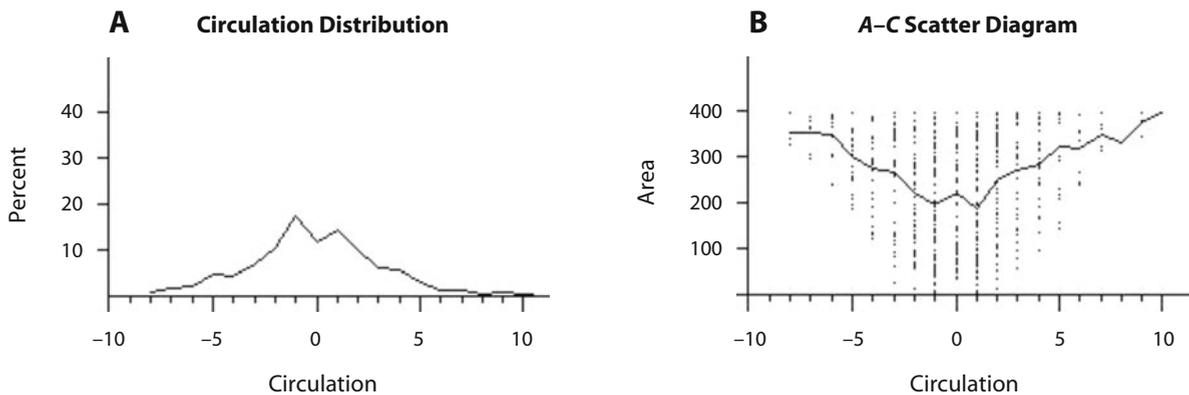


Figure 6. Simulated circulation C distribution (A) and its scatterplot with the area measure A (B) for older adults searching 48-distractor arrays. The corresponding observed C distribution and A – C scatterplot are shown in Figures 5A and 5B in Keech and Resca (2010).

Table 3
Changes of Simulation Results Caused by Variations in Model Parameters
for 48-Distractor Arrays Searched by the Older Adults

Set Group	Area of Conspicuity Extension	$g(r,\phi)$	R_I	Average Number of Saccades (SD)	Average Saccade Amplitude (deg) (SD)	Average Area Measure (deg ²) (SD)	Failure Rate (%)
1	1.0	enabled	0.05	8.80	6.92	285	9.0
				(4.97)	(4.08)	(130)	
				9.25	6.93	299	9.4
				(4.93)	(4.20)	(122)	
				9.36	6.96	298	7.3
				(4.92)	(4.09)	(123)	
2	1.1	enabled	0.05	6.94	7.47	259	0.3
				(4.18)	(4.27)	(125)	
				7.31	7.42	268	0
				(4.15)	(4.21)	(126)	
				7.14	7.40	269	0
				(4.13)	(4.19)	(127)	
3	1.2	enabled	0.05	6.59	7.53	256	0
				(3.86)	(4.23)	(128)	
				6.50	7.64	260	0
				(3.72)	(4.33)	(123)	
				6.69	7.54	264	0
				(3.81)	(4.15)	(124)	
4	1.3	enabled	0.05	5.83	7.63	244	0
				(3.45)	(4.17)	(122)	
				5.80	7.64	245	0
				(3.41)	(4.25)	(121)	
				6.13	7.55	250	0
				(3.62)	(4.18)	(126)	
5	1.4	enabled	0.05	5.65	7.96	253	0
				(3.42)	(4.30)	(126)	
				5.53	7.67	243	0
				(3.17)	(4.16)	(121)	
				5.68	7.75	255	0
				(3.41)	(4.18)	(124)	
6	1.2	disabled	0.05	6.60	8.05	236	0
				(3.90)	(4.42)	(117)	
				6.53	8.12	256	0
				(3.79)	(4.41)	(122)	
				6.40	7.98	259	0
				(3.64)	(4.44)	(123)	
7	1.2	enabled	1	7.13	7.37	258	3.8
				(4.71)	(4.05)	(128)	
				7.11	7.34	257	3.7
				(4.65)	(4.03)	(125)	
				7.36	7.34	264	4.1
				(4.61)	(3.94)	(125)	
8	1.2	enabled	5	7.63	6.87	264	32.6
				(5.03)	(3.56)	(126)	
				7.76	6.90	263	32.6
				(5.03)	(3.59)	(127)	
				7.71	6.93	267	34.1
				(4.95)	(3.60)	(126)	

Note—Shaded row shows the results of three independent 630-simulation sets where all the model parameters have values that optimize matching to all the measured data.

cient increases from 1.0 to 1.4 in the first five main rows in Table 3, the average saccade amplitude increases significantly (by about 11%), whereas the average number of saccades needed to capture the target decreases even more

substantially (by about 62%). Similarly, the average area measure A decreases by about 18% as the AC extension coefficient is increased from 1.0 to 1.4. The failure rate rises to about 9% from the baseline value around 0 when the

AC extension coefficient is reduced to 1.0 from the baseline value of 1.2. Upon reflection, all these results follow precisely the trends that one should expect. On the other hand, these results also indicate that the performance of the computational model depends rather sensitively on the selection and parameterization of the AC and its radius.

Set Group 6 corresponds to three 630-simulation sets for which the angle shift factor is disabled—that is, $g(r, \phi) = 1$, meaning that no specific bias to saccading consecutively in a forward direction is assumed (contradicting experimental evidence). Comparing Group Set 6 with the baseline, the average saccade amplitude is further increased by about 6%, whereas the average number of saccades needed to capture the target and the average area measure A remain virtually unchanged. The saccade direction cosine autocorrelation with the angle shift factor disabled is further shown in Figure 7. Its comparison with Figure 5, corresponding to the baseline, confirms that the bias toward consecutively saccading forwardly has indeed been removed, since even $R(l = 1)$ has become negative in Figure 7. So, there is always prevalence of 90° – 270° turning of direction for all intervening saccades in Figure 7, as a result of reversals at the display boundary, unlike in the baseline case, where the forward bias in consecutive saccades overcomes that effect for $l = 1$. Thus, as was expected, the fidelity of the simulated saccade direction cosine autocorrelation is most compromised when the angle shift factor is disabled, as compared with the observed autocorrelation, shown in Figure 1C in Keech and Resca (2010) on an expanded scale.

In Set Groups 7 and 8, the R_l coefficient is increased from the baseline value of 0.05 to 1.0 and 5.0 values, respectively. For the 48-distractor arrays, we have $n = 1$ in Equation 4; hence, $h(l) = R_l / (l - 1)$, where l is the number of intervening saccades.¹³ Increasing R_l to 1.0 and 5.0 produces a substantial increase in the average number of saccades of approximately 9% and 17%, respectively, and a corresponding decrease in the average saccade amplitude of 3% and 10%, respectively. Increasing R_l corresponds to decreasing memory or even favoring returns, which evidently results in a greater number of useless saccades to previously inspected locations (assuming that there was nearly perfect identification of all the search items that fell within each previous AC, which is not exactly the case but

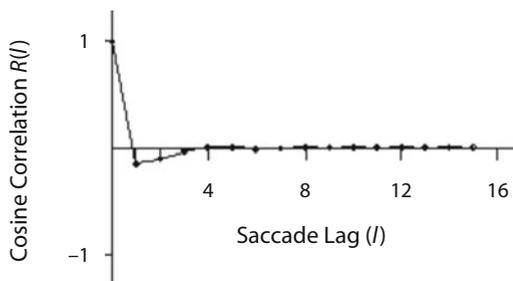


Figure 7. Simulated saccade direction cosine autocorrelation for 48-distractor arrays searched by older adults, with the angle shift factor disabled. Compare with Figure 5 and with the corresponding observed correlation, shown in Figure 1C in Keech and Resca (2010) on an expanded scale.

still a reasonably good approximation of our model). In turn, the failure rate is expected to substantially increase, since the target may not be detected within the maximum number (18) of allowed saccades. Indeed, the failure rate, which is 0% for the baseline case, increases to about 4% for $R_l = 1$ and jumps to about 33% for $R_l = 5$.

From all these observations and comparisons, we may conclude that the AC extension coefficient, the R_l coefficient, and the angle shift factor are all critical, and independently so, for the model's ability to faithfully reproduce measured results. Still, the most essential parameter is the first one, followed by the other two, in that order.

“Memoryless” Simulations

We recall that setting $h(l) = 1$ means that after any number l of intervening saccades, there is no memory of objects in the corresponding AC fixated originally. Conversely, either $h(l) < 1$ or $h(l) > 1$ implies that after l intervening saccades, memory is used to either inhibit or favor the chance of returning to objects in the corresponding AC fixated originally. Since we wish to establish whether memory operates critically in the visual search process under investigation, we have generated and analyzed a series of “memoryless” computer simulations, by setting $h(l) = 1$ identically for all $l = 2, 3, 4, \dots$

Figure 8 shows the return rate of a memoryless simulation, which should be compared with the corresponding observations for young adults searching arrays with 48 distractors, shown as a thick line in Figure 19C in Keech et al. (2009) and superimposed as a thick line in the present Figure 8 as well. The disparity between the memoryless and the observed return rates is both qualitatively and quantitatively overwhelming, with the former exceeding the latter by a factor of 5 or more, depending on the delay. Likewise, Figure 9 shows the frequency distribution of the number of saccades needed to detect the target for memoryless computer simulations (panel A) and the corresponding observations for young adults searching arrays with 48 distractors

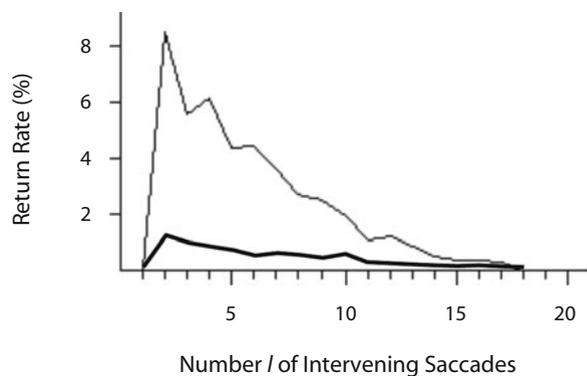


Figure 8. Memoryless simulation of the return rate (percentage) to an original fixation location after l intervening saccades, corresponding to young adults searching arrays with 48 distractors. The observed return rate is shown as a thick line in Figure 19C in Keech, Greenwood, Resca, and Parasuraman (2009) and is superimposed as a thick line in the present Figure 8 for comparison.

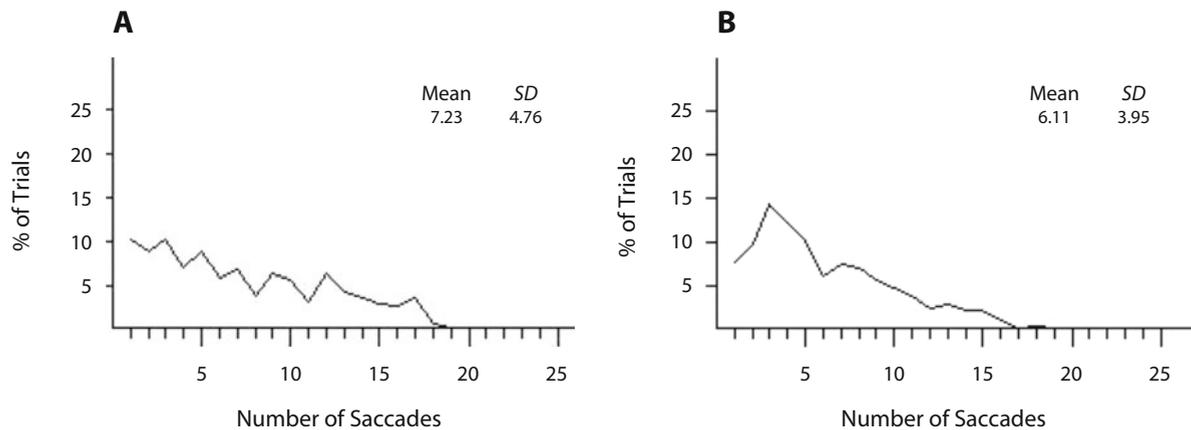


Figure 9. Frequency distributions (percentages of trials) of the number of saccades per trial for young adults searching arrays with 48 distractors: (A) Memoryless simulations; (B) actual observations.

(panel B). The memoryless search typically requires many more saccades per trial than observed, resulting in a much further stretched distribution: Its mean of 7.23 saccades has been stretched by 18%, and its standard deviation of 4.76 saccades has also increased by 21%, as compared with the observed saccade number distribution. Correspondingly, the failure rate for detecting the target jumps to about 33% for the memoryless simulations, as compared with 0.9% for the actual observations (cf. Table 2).

Similar major discrepancies for memoryless simulations are obtained for either young or older adults searching arrays of all sizes—and the larger ones, in particular. Substantial discrepancies remain even if we match the observed return rate for immediate return—that is, if we reinstate in the simulations a correspondingly optimized $h(l)$ for $l = 2$ exclusively. That represents visual search with only the “shortest memory,” which was already noted in Figure 7B in Motter and Belky (1998a). Still, return rates for shortest memory simulations greatly exceed the observed return rates for subsequent ($l > 2$) numbers of intervening saccades.

We have further verified that no change in other factors or parameters in our computational model—for example, the angle shift factor or the AC extension coefficient—can compensate in any way for the major failures of either the memoryless or the shortest memory computer simulations. This demonstrates that the memory factor is critical to our computational model and simulations—hence, that a corresponding long-term, although possibly coarse, form of memory definitely supports the visual search task that we investigated experimentally (Keech et al., 2009).

DISCUSSION

The main purpose of our computational model and simulations was to confirm or reject the Motter–Belky model, augmented with some form of long-range memory, if necessary. We did that by translating the basic points of such a source model into a precise mathematical and probabilistic formulation, whose computational implementation and simulations we then compared with our observations

of corresponding eye movements and fixations (Keech, 2006; Keech et al., 2009; Keech & Resca, 2010). We expected that to deepen our understanding of three basic psychoneural mechanisms (AC, forward bias, long-range memory) and two task-specific geometric-spatial constraints on the eye trajectories (central start and display confinement¹⁴) that appear to underlie and regulate the conjunctive visual search task that we investigated.

Although the *ideal* of the model would be to generate computer simulations that mimic in every respect the observed eye movements of an average or ideal viewer, that would be asking too much, of course. We thus identified some most salient features and characteristics of the eye trajectories and some corresponding measures that should be optimally reproduced in the simulated trajectories, on average. Those include autocorrelation functions of the saccade direction cosines and scatterplots of certain measures of circulation and subtended area that we previously introduced for the analysis and characterization of the observed eye trajectories (Keech, 2006; Keech & Resca, 2010). We further considered saccade number distributions, failure rates to detect the target, and saccade amplitude distributions as critical benchmarks of performance for our simulations.

The results of our simulations and their comparison with the observed eye movements generally support the Motter–Belky model, provided that it is adequately augmented by a form of long-range memory. More specifically, salient measures and benchmark comparisons indicate that our computational model can be consistently parameterized to simulate observations to the following extent. (1) The saccade direction cosine autocorrelations derived from the simulations follow the same trend with increasing array size as the saccade direction cosine autocorrelations obtained from the measurements, confirming, in particular, the competing effects of the forward bias and of the display boundaries on the progression of the eye movement trajectories. (2) The distributions of the circulation C derived from the simulations show the same bimodal characteristics as those derived from the observations, both indicating the presence of spiraling or circulat-

ing trajectories. (3) The scatterplots of the area measure A and circulation C display consistent curvilinear regressions for both the simulated and the measured eye trajectories, which also provides a signature of spiraling or circulating trajectories (Keech & Resca, 2010). (4) The distributions of the number of saccades and the failure rates for target capture derived from the simulations reasonably match the measured number distributions and failure rates for all array sizes. (5) Vice versa, the saccade amplitude distributions emerging from the simulations *cannot* always consistently reproduce the measured distributions.

The outcome mentioned in (5) is clearly at odds with all the previous ones (1–4), indicating a difficulty or limitation in our computational model, which has the following origin and price. Unlike in most other models of saccade generation, we wished to take into account the fact that items at different distances from the current fixation have different probabilities of being selected for the ensuing saccade. Indeed, according to Point 2 of the Motter–Belky model, if the target is not detected within the current AC, a saccade is made most likely to an item just beyond that AC. We have thus assumed a specific distance factor $f(r)$ that is meant to contribute, to the probability $p_{i \rightarrow j}$ of saccading from object to object, the “a priori” probability of saccading to a point at a distance r from the current fixation. We know from empirical evidence, subsumed in Point 2 of the Motter–Belky model, that such “a priori” probability is somehow related to the AC, but we have no way of knowing precisely how. So, we can only turn to the measured SAD to determine $f(r)$ empirically, by simply dividing that SAD by the ISD distribution (see Appendix B). For the very first saccade from the central fixation, that works just fine, since only $f(r)$ contributes to $p_{i \rightarrow j}$ in Equation 2. However, for the following saccades, the angle shift $g(r, \phi)$ and memory $h(l)$ factors also contribute to $p_{i \rightarrow j}$ in Equation 2. Then, the measured SAD no longer reflects exclusively the “a priori” distance probability but, rather, some aggregate of all psychoneural mechanisms (and task-specific constraints) that are involved in the search. That interaction limits the attainment of complete self-consistency by the model when it tries to reproduce “a posteriori,” through the computer simulations, the measured SAD that it assumed “a priori.”

As was noted in the Computer Simulations section, just setting $h(l) = 1$ in the simulations allows recovery of the measured SAD self-consistently. However, far greater discrepancies with other observations—and return rates, in particular—arise for those memoryless simulations. In principle, one could solve the problem by initially guessing $f(r)$ —and its range, in particular—more or less independently of the SAD and then by adjusting $f(r)$ iteratively until the observed SAD and everything else is ultimately reproduced. On the other hand, that could quickly escalate into a cumbersome and rather pointless curve-fitting exercise, undermining the original intent of relative simplicity and transparency of the model. Until a more fundamental understanding (and a corresponding modeling) of the relation between the AC and SAD is attained, we may as well accept provisionally this current limitation of our model.

On the other hand, we should also point out that we were able to understand more deeply the connection and discover the internal consistency between the observed SAD and the average number of saccades per trial. That is essentially the result of the intertwined roles that the AC plays in both the identification of objects and their memory tagging to avoid reinspection. In just one sentence, one cannot *remember* as a nontarget what has not first been *identified* as a nontarget, whether correctly or incorrectly. We present in Appendix C such a revealing and independent analytical formulation and statistical estimate of that connection.

Furthermore, despite the difficulty in consistently reproducing measured SADs and the crude assumption of an equipotent AC, Point 2 of the Motter–Belky model still seems to be supported by our computer simulations, at least qualitatively. Indeed, in units of c-ANND, the most probable saccade amplitudes for young adults searching arrays with 12, 24, 48, or 96 distractors are 1.0, 1.4, 1.8, and 1.9, respectively, whereas those for the older adults are 0.8, 1.1, 1.1, and 1.4, respectively (cf. Figure 8 in Keech et al., 2009). If saccades are most likely made to items just beyond the current AC, those most probable amplitudes should roughly correspond to the extension coefficients that multiply the c-ANND to yield the AC radius, by definition. The extension coefficients reported in Table 2 for the optimally parameterized simulations seem to confirm that expectation, at least qualitatively.

However, Figure 5B in Motter and Belky (1998b) shows a remarkable scaling property of the sensitivity curves corresponding to the AC, which become essentially independent of array size once distances are expressed in ANND units. That would suggest that the extension coefficient should always remain relatively close to one and independent of array size, which is not what our model and simulations predict, as is reported in Table 2. Our crude assumption of an *equipotent* AC may allow such apparent inconsistency with the results just reported in the preceding paragraph.

Finally, it may be enlightening to further compare our computational model and results with those of (1) an alternative approach to active visual search based on time series analysis and complex dynamics (Aks et al., 2002), (2) a previous set of models and simulations asserting that visual search has memory (Peterson et al., 2001), and (3) a contrasting perspective of an ideal Bayesian observer (Najemnik & Geisler, 2005). Since these comparisons require fairly detailed reviews and analyses of corresponding articles, they are reported in the three separate subsections of Supplemental Appendix A, posted online.

CONCLUSION

We have demonstrated through extensive computer simulations that a model of conjunctive visual search derived originally by Motter and his collaborators from many years of insightful studies and careful observations “does what we say it does” (to use Wolfe’s phrase) with its basic psychoneural mechanisms—that is, AC and forward bias. However, that model can account quantitatively for

observations such as return rates only if it is augmented by some form of long-range memory. The precise nature of that form of memory stands to be further investigated and determined. The possibility of a high-capacity but low-resolution *memory for the search path*, suggested by Dickinson and Zelinsky (2007), is compatible with our results and particularly appealing, but other alternatives are also possible (Pashler, 1987; Peterson, Beck, & Vomela, 2007; Peterson et al., 2001). On the other hand, ideal Bayesian searcher simulations by Najemnik and Geisler (2005) suggest that “all that is actually needed is a coarse memory sufficient to support the inhibition of return” (p. 390) at *short range*. We have attributed the origin of this apparent discrepancy to the profoundly different nature of the displays presented to our subjects and to those who participated in the experiments of Najemnik and Geisler (see Supplemental Appendix A).

Horowitz and Wolfe (2001) have previously suggested the following:

The visual system appears to perform adequately without relying on a rich memory representation. The attentional system acts like a “cognitive miser” (S. E. Taylor, 1981), preferring rapid, adequate performance to slow, accurate performance. This is consistent with an emerging body of evidence concerning the transient nature of visual representations. Recent research suggests that, instead of building up a rich, detailed visual representation, the visual system quickly and efficiently constructs a brief sketch of the visual world, which is discarded in the next instant. . . . If vision in general consists of a series of ephemeral and sparse descriptions of the world, then it becomes less surprising that visual search is not accomplished by building up a representation of the attended areas of the scene. A visual system that uses the world as its memory (O’Regan, 1992) would allow the world itself (mediated, perhaps, by a salience map; Koch & Ullman, 1985) to direct attention, rather than maintaining a map of inhibited locations. . . . Since the data indicate no savings from one target to the next, we conclude that the visual system does not rely on inhibitory tagging to guide attention during search. This conclusion agrees with our previous experiments (Horowitz & Wolfe, 1998), as well as with the emerging view of “just in time vision.” While memory is certainly crucial to perform various components of this search task, it is not employed to improve attentional “foraging.” (p. 282)

All that may be true, but it may not be the whole story. Perhaps the “attentional system” is a “virtual memorist” that often acts or looks like a “cognitive miser.” Namely, the visual system may not use memory (possibly long-ranged and demanding) when that is not essential to improve search efficiency—for example, if it is possible to use “the world as its memory.” That may very well happen most of the time, but not necessarily all the time. In our displays, for example, with many completely randomized and *uninformative* items (except for

color), it is just impossible for the visual system to use “the world as its memory.” For our visual search task (and many similar ones routinely performed in laboratories), the one and only critical element that can make the search drastically more efficient is to use some form of long-range memory. Evidence from a *survivor function* approach (Keech et al., 2009), *power spectra* (Keech & Resca, 2010), and the computational model presented in this article, combined with similar evidence provided by other authors (Dickinson & Zelinsky, 2007; Peterson et al., 2001), leaves no doubt that the eye–brain system is quite capable of resorting to long-range memory, when that is critically needed.¹⁵

Surely, the eye–brain system is unimaginably “smart” and “adaptable” in carrying out most efficiently the multitude of extremely different search tasks that it confronts daily. It has certainly developed a broad range of psychoneural resources and search strategies that it can bring to bear most effectively for whatever and whenever is needed. There is no single experiment or condition that can possibly tell us all that the eye–brain system can or cannot do, or how it does it “normally.” Rather, we may have to accept some sort of an “uncertainty principle”: *What we measure always depends on how we measure it.*

Last, but not least, we have developed a complementary approach to determine the average number of saccades per trial and its relation to the SAD from a basic understanding of the combined roles that the AC plays in attention and memory. We report that method in a separate Appendix C.

AUTHOR NOTE

The authors are greatly indebted to Pamela Greenwood and Raja Parasuraman for introducing and training them in this research area and project. The authors are also greatly indebted to Brad Motter for countless and extremely valuable comments and suggestions throughout this entire work. It is with great sorrow that L.R. acknowledges the loss of T.D.K., his former student and dear colleague. L.R. is particularly grateful to Deborah Aks, Bradley Gibson, Todd Horowitz, and Matthew Peterson for their thorough reviews and critical contributions to both substance and presentation of this article. Although differing in their approaches, those three reviewers and the Associate Editor have been equally and extremely effective in improving systematically the authors’ input. So, this article ultimately reflects the remarkable expertise collectively provided to the authors by all the individuals mentioned above. Technical assistance by Richard Cecil is also gratefully acknowledged. Correspondence concerning this article should be addressed to L. Resca, Physics Department, Catholic University of America, Washington, DC 20064 (e-mail: resca@cua.edu).

REFERENCES

- AKS, D. J. (2005). *1/f* dynamic in complex visual search: Evidence for self-organized criticality in human perception. In M. A. Riley & G. C. Van Orden (Eds.), *Tutorials in contemporary nonlinear methods for the behavioral sciences* (pp. 319-352). Retrieved in 2009 from www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp.
- AKS, D. J. (2009). Studying temporal and spatial patterns in perceptual behavior: Implications for dynamical structure. In S. Guastello, M. Koopmans, & D. Pincus (Eds.), *Chaos and complexity in psychology: The theory of nonlinear dynamical systems* (pp. 132-176). New York: Cambridge University Press.
- AKS, D. J., ZELINSKY, G., & SPOTT, J. C. (2002). Memory across eye-movements: *1/f* dynamic in visual search. *Nonlinear Dynamics, Psychology, & Life Sciences*, *6*, 1-25.

- ARANI, T., KARWAN, M. H., & DRURY, C. G. (1984). A variable-memory model of visual search. *Human Factors*, **26**, 631-639.
- BAK, P., TANG, C., & WIESENFELD, K. (1987). Self-organized criticality: An explanation of $1/f$ noise. *Physical Review Letters*, **59**, 381-384.
- BALL, K. K., BEARD, B. L., ROENKER, D. L., MILLER, R. L., & GRIGGS, D. S. (1988). Age and visual search: Expanding the useful field of view. *Journal of the Optical Society of America*, **5**, 2210-2219.
- BALLARD, D. H., HAYHOE, M. M., POOK, P. K., & RAO, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral & Brain Sciences*, **20**, 723-767.
- BOCCIGNONE, G., & FERRARO, M. (2004). Modeling gaze shifts as a constrained random walk. *Physica A*, **331**, 207-218.
- BROCKMANN, D., & GEISEL, T. (2000). The ecology of gaze shifts. *Neurocomputing*, **32**, 643-650.
- CANOSA, R. L. (2005). Modeling selective perception of complex, natural scenes. *International Journal of Artificial Intelligence Tools: Architectures, Languages, Algorithms*, **14**, 233-260.
- CANOSA, R. L. (2009). Real-world vision: Selective perception and task. *ACM Transactions on Applied Perception*, **6**, Art. 11, 1-34.
- CAVE, K. R., & WOLFE, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology*, **22**, 225-271.
- CHUN, M. M., & WOLFE, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology*, **30**, 39-78.
- COHEN, E., & RUPPIN, E. (1999). From parallel to serial processing: A computational study of visual search. *Perception & Psychophysics*, **61**, 1449-1461.
- COUSINEAU, D., & SHIFFRIN, R. M. (2004). Termination of a visual search with large display-size effects. *Spatial Vision*, **17**, 327-352.
- DANZIGER, S., KINGSTONE, A., & SNYDER, J. J. (1998). Inhibition of return to successively stimulated locations in a sequential visual search paradigm. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1467-1475.
- DICKINSON, C. A., & ZELINSKY, G. J. (2007). Memory for the search path: Evidence for a high-capacity representation of search history. *Vision Research*, **47**, 1745-1755.
- DUNCAN, J., & HUMPHREYS, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, **96**, 433-458.
- DUNCAN, J., & HUMPHREYS, G. W. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 578-588.
- EGETH, H. E., VIRZI, R. A., & GARBART, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 32-39.
- ENGEL, F. L. (1971). Visual conspicuity, directed attention and retinal locus. *Vision Research*, **11**, 563-576.
- ENGEL, F. L. (1974). Visual conspicuity and selective background interference in eccentric vision. *Vision Research*, **14**, 459-471.
- ENGEL, F. L. (1977). Visual conspicuity, visual search and fixation tendencies of the eye. *Vision Research*, **17**, 95-108.
- FECTEAU, J. H., & MUNOZ, D. P. (2006). Saliency, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences*, **10**, 382-390.
- FINDLAY, J. M., & BROWN, V. (2006). Eye scanning of multi-element displays: II. Saccade planning. *Vision Research*, **46**, 216-227.
- FINDLAY, J. M., BROWN, V., & GILCHRIST, I. D. (2001). Saccade target selection in visual search: The effect of information from the previous fixation. *Vision Research*, **41**, 87-95.
- GIBSON, B. S., & EGETH, H. (1994). Inhibition of return to object-based and environment-based locations. *Perception & Psychophysics*, **55**, 323-339.
- GILCHRIST, I. D., & HARVEY, M. (2000). Refixation frequency and memory mechanisms in visual search. *Current Biology*, **10**, 1209-1212.
- GODIEN, R., & THEEUWES, J. (2002). Programming of endogenous and exogenous saccades: Evidence for a competitive integration model. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 1039-1054.
- 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.
- HENDERSON, J. M. (1992). Visual attention and eye-movement control during reading and picture viewing. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 260-283). New York: Springer.
- HOOG, I. T. C., & ERKELENS, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics*, **58**, 969-976.
- HOOG, I. T. C., & ERKELENS, C. J. (1998). Adjustment of fixation duration in visual search. *Vision Research*, **38**, 1295-1302.
- HOROWITZ, T. S., & WOLFE, J. M. (1998). 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., ALVAREZ, G. A., COHEN, M. A., & KUZMOVA, Y. I. (2009). The speed of free will. *Quarterly Journal of Experimental Psychology*, **62**, 2262-2288.
- HUMPHREYS, G. W., & MÜLLER, H. J. (1993). Search via recursive rejection (SERR): A connectionist model of visual search. *Cognitive Psychology*, **25**, 43-110.
- ITTI, L., & KOCH, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, **40**, 1489-1506.
- ITTI, L., & KOCH, C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience*, **2**, 1-11.
- ITTI, L., KOCH, C., & NIEBUR, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **20**, 1254-1259.
- KEECH, T. D. (2006). *Dynamics of spontaneous saccades in a conjunctive visual search task*. Unpublished doctoral thesis, Catholic University of America. Available at <http://libraries.cua.edu/welcome.html> under the Dissertations from CUA database selection.
- KEECH, T. D., GREENWOOD, P. M., RESCA, L., & PARASURAMAN, R. (2009). *How do eyes and brain search a randomly structured uninformative scene? Exploiting a basic interplay of attention and memory*. Unpublished manuscript.
- KEECH, T. D., & RESCA, L. (2010). Eye movement trajectories in active visual search: Contributions of attention, memory, and scene boundaries to pattern formation. *Attention, Perception, & Psychophysics*, **72**, 114-141.
- KLEIN, R. [M.] (1988). Inhibitory tagging system facilitates visual search. *Nature*, **334**, 430-431.
- KLEIN, R. M., & MACINNIS, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science*, **10**, 346-352.
- KLEIN, R. M., & TAYLOR, T. L. (1994). Categories of cognitive inhibition with reference to attention. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 113-150). San Diego: Academic Press.
- KOCH, C., & ULLMAN, S. (1985). Shifts in selective visual attention: Toward the underlying neural circuitry. *Human Neurobiology*, **4**, 219-227.
- KRAISS, K. F., & KNÄEUPER, A. (1982). Using visual lobe area measurements to predict visual search performance. *Human Factors*, **24**, 673-682.
- KWOK, T., & SMITH, K. A. (2005). Optimization via intermittency with a self-organizing neural network. *Neural Computation*, **17**, 2454-2481.
- LAND, M. L., MENNIE, N., & RUSTED, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, **28**, 1311-1328.
- LE MEUR, O., LE CALLET, P., BARBA, D., & THOREAU, D. (2006). A coherent computational approach to model bottom-up visual attention. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **28**, 802-817.
- MCCARLEY, J. S., WANG, R. F., KRAMER, A. F., IRWIN, D. E., & PETERSON, M. S. (2003). How much memory does oculomotor search have? *Psychological Science*, **14**, 422-426.
- MCPPEEK, R. M., SKAVENSKI, A. A., & NAKAYAMA, K. (2000). Concurrent processing of saccades in visual search. *Vision Research*, **40**, 2499-2516.
- MOTTER, B. C., & BELKY, E. J. (1998a). The guidance of eye movements during active visual search. *Vision Research*, **38**, 1805-1815.
- MOTTER, B. C., & BELKY, E. J. (1998b). The zone of focal attention during active visual search. *Vision Research*, **38**, 1007-1022.
- MOTTER, B. C., & HOLSAPPLE, J. W. (2000). Cortical image density determines the probability of target discovery during active search. *Vision Research*, **40**, 1311-1322.

- MOTTER, B. C., & HOLSAPPLE, J. W. (2001). Separating attention from chance in active visual search. In J. Braun, C. Koch, & J. Davis (Eds.), *Visual attention and neural circuits* (pp. 159-175). Cambridge, MA: MIT Press.
- MOTTER, B. C., & HOLSAPPLE, J. W. (2007). Saccades and covert shifts of attention during active visual search: Spatial distributions, memory, and items per fixation. *Vision Research*, **47**, 1261-1281.
- MOTTER, B. C., & SIMONI, D. A. (2007). The roles of cortical image separation and size in active visual search performance. *Journal of Vision*, **7**(2), 1-15. doi:10.1167/7.2.6
- MOTTER, B. C., & SIMONI, D. A. (2008). Changes in the functional visual field during search with and without eye movements. *Vision Research*, **48**, 2382-2393.
- MÜLLER, H. J., & VON MÜHLENEN, A. (2000). Probing distractor inhibition in visual search: Inhibition of return. *Journal of Experimental Psychology: Human Perception & Performance*, **32**, 1591-1605.
- NAJEMNIK, J., & GEISLER, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, **434**, 387-391.
- NAVALPAKKAM, V., & ITTI, L. (2002). A goal oriented attention guidance model. In *Lecture Notes in Computer Science: Proceedings of the Second International Workshop on Biologically Motivated Computer Vision* (Vol. 2525, pp. 453-461). London: Springer.
- NAVALPAKKAM, V., & ITTI, L. (2006a). An integrated model of top-down and bottom-up attention for optimizing detection speed. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Vol. 2, pp. 2049-2056). Los Alamitos, CA: IEEE Computer Society Press.
- NAVALPAKKAM, V., & ITTI, L. (2006b). Top-down attention selection is fine grained. *Journal of Vision*, **6**(11), 1180-1193.
- NAVALPAKKAM, V., & ITTI, L. (2007). Search goal tunes visual features optimally. *Neuron*, **53**, 605-617.
- NOTON, D., & STARK, L. W. (1971). Scanpaths in eye movements during pattern perception. *Science*, **171**, 308-311.
- O'REGAN, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, **46**, 461-488.
- PACZUSKI, M., MASLOV, S., & BAK, P. (1996). Avalanche dynamics in evolution, growth, and depinning models. *Physical Review E*, **53**, 414-443.
- PARKHURST, D., LAW, K., & NIEBUR, E. (2002). Modeling the role of saliency in the allocation of overt visual attention. *Vision Research*, **42**, 107-123.
- PASHLER, H. (1987). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception & Psychophysics*, **41**, 191-201.
- PELZ, J. B., & CANOSA, R. L. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research*, **41**, 3587-3596.
- PELZ, J. B., CANOSA, R. L., & BABCOCK, J. (2000). Extended tasks elicit complex eye movement patterns. In *Proceedings of the ACM SIGCHI Eyetracking Research and Applications Symposium* (pp. 37-43). New York: ACM Press.
- PETERS, R. J., & ITTI, L. (2007a). Applying computational tools to predict gaze direction in interactive visual environments. *ACM Transactions on Applied Perception*, **5**, Art. 9, 1-21.
- PETERS, R. J., & ITTI, L. (2007b). Beyond bottom-up: Incorporating task-dependent influences into a computational model of spatial attention. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-8). Los Alamitos, CA: IEEE Computer Society Press.
- PETERS, R. J., IYER, A., ITTI, L., & KOCH, C. (2005). Components of bottom-up gaze allocation in natural images. *Vision Research*, **45**, 2397-2416.
- PETERSON, M. S., BECK, M. R., & VOMELA, M. (2007). Visual search is guided by prospective and retrospective memory. *Perception & Psychophysics*, **69**, 123-135.
- PETERSON, M. S., KRAMER, A. F., WANG, R. F., IRWIN, D. E., & MCCARLEY, J. S. (2001). Visual search has memory. *Psychological Science*, **12**, 287-292.
- POMPLUN, M., CARBONE, E., KOESLING, H., SICHELSCHMIDT, L., & RITTER, H. (2006). Computational models of visual tagging. In G. Rickheit & I. Wachsmuth (Eds.), *Situated communications* (pp. 213-246). Berlin: Mouton de Gruyter.
- POMPLUN, M., REINGOLD, E. M., & SHEN, J. (2003). Area activation: A computational model of saccadic selectivity in visual search. *Cognitive Science*, **27**, 299-312.
- PRIVITERA, C., & STARK, L. W. (2000). Algorithms for defining visual regions-of-interest: Comparisons with eye fixations. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **22**, 970-982.
- 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.
- SANDERS, A. F. (1970). Some aspects of the selective process in the functional visual field. *Ergonomics*, **13**, 101-117.
- SCINTO, L. F. M., PILLALAMARRI, R., & KARSH, R. (1986). Cognitive strategies for visual search. *Acta Psychologica*, **62**, 263-292.
- SNYDER, J. J., & KINGSTONE, A. (2000). Inhibition of return and visual search: How many separate loci are inhibited? *Perception & Psychophysics*, **62**, 452-458.
- SPALEK, T. M., & HAMMAD, S. (2004). Supporting the attentional momentum view of IOR: Is attention biased to go right? *Perception & Psychophysics*, **66**, 219-233.
- SPROTT, J. C. (2003). *Chaos and time-series analyses*. New York: Oxford University Press.
- STEPHEN, D. G., MIRMAN, D., MAGNUSON, J. S., & DIXON, J. A. (2009). Lévy-like diffusion in eye movements during spoken-language comprehension. *Physical Review E*, **79**, 056114_1-6.
- TAKEDA, Y., & YAGI, A. (2000). Inhibitory tagging in visual search can be found if search stimuli remain visible. *Perception & Psychophysics*, **62**, 927-934.
- TAYLOR, S. E. (1981). A categorization approach to stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 88-114). Hillsdale, NJ: Erlbaum.
- TIPPER, S. P., WEAVER, B., JERREAT, L. M., & BURAK, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 478-499.
- TORRALBA, A. (2003). Modeling global scene factors in attention. *Journal of the Optical Society of America*, **20**, 1407-1418.
- TORRALBA, A., OLIVA, A., CASTELHANO, M. S., & HENDERSON, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features on object search. *Psychological Review*, **113**, 766-786.
- TREISMAN, A. M., & GELADE, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97-136.
- TREISMAN, A. [M.], & GORMICAN, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, **95**, 15-48.
- TREISMAN, A. [M.], & SATO, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 459-478.
- TSOTSOS, J. K., CULHANE, S. M., WAI, W. Y. K., LAI, Y., DAVIS, N., & NUFLO, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence*, **78**, 507-545.
- USHER, M., STEMLER, M., & OLAMI, Z. (1995). Dynamic pattern formation leads to 1/f noise in neural populations. *Physical Review Letters*, **74**, 326-329.
- WIDDEL, H. (1983). A method of measuring the visual lobe area. In R. Groner, C. Menz, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and psychological functions: International views* (pp. 73-83). Hillsdale, NJ: Erlbaum.
- WILLIAMS, L. G. (1966). Target conspicuity and visual search. *Human Factors*, **8**, 80-92.
- WOLFE, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, **1**, 202-238.
- WOLFE, J. M., BUTCHER, S. J., LEE, C., & HYLE, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 483-502.
- WOLFE, J. M., FRIEDMAN-HILL, S. R., STEWART, M. I., & O'CONNELL, K. M. (1992). The role of categorization in visual search for orientation. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 34-49.
- WOLFE, J. M., HOROWITZ, T. S., KENNER, N., HYLE, M., & VASAN, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*, **44**, 1411-1426.
- WOLFE, J. M., & POKORNY, C. W. (1990). Inhibitory tagging in vi-

sual search: A failure to replicate. *Perception & Psychophysics*, **48**, 357-362.

ZELINSKY, G. J., & SHEINBERG, D. L. (1997). Eye movements during parallel-serial visual search. *Journal of Experimental Psychology*, **23**, 244-262.

NOTES

1. For example, Aks (2009) reminds us that "a critical point about the most common use of RTs in cognitive research is this: These reflect distributions of manually recorded 'finishing times,' and thus miss the time course leading up to the final response. Because they are a cumulative measure (a net combination of multiple implicit processes), RTs are by definition 'confounded'" (p. 135). Conversely, eye movements "(a) provide a real-time behavioral record, (b) they are highly correlated with attention, manual RTs, and various physiological measures, and (c) they operate at the interface of sensory-motor processes."

2. *Active visual search* is performed with eye movements, or *overt shifts* of attention, as distinguished from *covert visual search*, in which it is possible or required to maintain eye fixation at a given location while searching elsewhere through *covert shifts* of attention. This is the terminology that we adopt here, which should not be confused with that of *active vision* used in another context (Canosa, 2005, 2009).

3. For further review and investigation of the interplay between bottom-up and top-down factors in visual search guidance, see Wolfe, Butcher, Lee, and Hyle (2003) and Wolfe, Horowitz, Kenner, Hyle, and Vasani (2004).

4. For example, the scatter diagram reported in Figure 8 in Itti and Koch (2000) between simulated and actual "reaction times" does not appear to produce a significant correlation.

5. One may further recall that this is not merely a perceptual problem of acuity threshold in peripheral vision (Motter & Simoni, 2008).

6. For a brief review and summary of research on inhibition of return, see a corresponding subsection in the introduction of our companion paper (Keech & Resca, 2010).

7. The forward bias that is observed in consecutive saccades (Keech et al., 2009; Motter & Belky, 1998a) may also have other origins (Aks, 2009; Spalek & Hammad, 2004). Our investigation does not determine the origin(s) of such forward bias; hence, our computational model accounts for it phenomenologically.

8. This time-equivalent constraint on the maximum number (18) of fixations and saccades allowed in our simulations, in which the target is always present, should not be confused with the altogether different issue of when a search without any time constraint may nevertheless terminate spontaneously when subjects decide that the target may be absent (Chun & Wolfe, 1996; Cousineau & Shiffrin, 2004). A time limit of 4 sec was imposed in our visual search task in order to stimulate fast and efficient shifts of attention. On the other hand, reduction or removal of any time constraint could have produced alternative valuable information, such as that just mentioned, or allowed further study and characterization of saccade number distributions, for example, which may obey power laws (Brockmann & Geisel, 2000). For such scale-free distributions, consideration of tails and outliers is essential, whereas those are curtailed by a time constraint on the search. Studies of complex dynamics in visual search tasks relatively similar to ours, but with reduced or removed time constraints, have been conducted by Aks (2005, 2009) and Aks et al. (2002). See Appendix A for a brief review.

Failure to detect the target within 18 fixations in a simulation is assumed to reflect failure to press a button within the time limit of 4 sec by a subject in the experiment (Keech et al., 2009). On the other hand, there were cases in which a subject signaled target presence when, in fact, the target did not appear to fall within any terminal AC during the search (see Figure 9C in Keech & Resca, 2010, for such an instance). Such cases likely represent false alarms. They appeared to be relatively rare, but we did not attempt to identify them computationally and study their statistics more precisely. Correspondingly, we assumed no chance of false alarms in our simulations, which cannot be exactly the case but should not significantly affect our main conclusions either.

9. This trend is at least partly a consequence of reversals in saccade directions that occur more frequently at the display boundary for larger saccade amplitudes (Keech et al., 2009; Keech & Resca, 2010).

10. Since ϕ is an angular variable, it may have been preferable to use a von Mises factor $\exp(\kappa \cos \phi)$ rather than our Gaussian factor

$\exp[-(\phi/\phi_s)^2]$. On the other hand, those two factors practically coincide for $\kappa \approx 1/\sigma^2 \equiv 2/\phi_s^2 = 2$, since they are both narrowly peaked within the two forward quadrants (where $-\pi/2 < \phi < \pi/2$).

11. Note, however, that this represents just a phenomenological description. For example, the reduction of the *recency effect* expressed by $h(l)$ is at least partly due to a small but consistent tendency for about 3%-4% of all saccades to return to the item just previously fixated, presumably for an attentional purpose of immediate verification, rather than because of lack of recent memory (Keech et al., 2009; Motter & Belky, 1998a, Figure 7B). Furthermore, the forward bias expressed by $g(r, \phi)$ generates itself some recency effect, or short-range memory, removing the need to account for a corresponding contribution in the $h(l)$ memory factor. See later discussion of this and other *interaction* effects among psychoneural and task-constraint mechanisms in the Discussion section.

An alternative approach to accounting for long-range memory or a *primacy effect* would be to first realize that, in our visual search task and setup, the eye trajectories already tend to spiral outward from the display center with a forward bias and then *further* reduce the probability of return to central locations inspected initially through some form of coarse memory for those already inspected regions, rather than for the items inspected specifically. Such more complex type of modeling, which we did not attempt, could represent more closely the idea of some form of high-capacity but low-resolution *memory for the search path* (Dickinson & Zelinsky, 2007) or some form of systematic searching and *prospective memory* (Peterson, Beck, & Vomela, 2007).

12. Ill-fated circumstances prevented us from completing some statistical tests that were planned for some distributions, such as chi-square tests between the simulated results shown in Figure 4 and the corresponding observations, or between the simulations and observations shown in Figures 9A and 9B. Some of our inferences thus remain based on our subjective pattern recognition. Nevertheless, we believe that all the major features that we have identified are in fact unmistakable, and none of our major conclusions would have minimally changed had we also been able to complete whatever standard statistical test. Of course, each reader can and should reach his or her own conclusions by comparing the same figures that we have observed and presented. Unfortunately, to say the least, the same circumstances also limit our presentation of return rates.

13. The number l of intervening saccades in the memory factor $h(l)$ exceeds by one unit the lag considered in the saccade autocorrelation $R(l)$, which is also labeled as l , unfortunately.

14. Saccades tend to target search items, and the oculomotor system perceives that there are none of those outside the display, thus preventing saccades from straying out of the display boundary, for the most part (Keech & Resca, 2010). Our computational model assumes all that as occurring exactly.

15. Horowitz, Wolfe, Alvarez, Cohen, and Kuzmova (2009) also suggested that possibility, but in a different context and for a different type of experiment. By comparing voluntary shifts of attention, associated with a search strategy precisely determined by a certain experimental condition, with automatic shifts of attention, driven by an alternative experimental condition, they found that voluntary shifts of attention may only proceed at a much slower pace than involuntary shifts. Since the voluntary shifts were associated with *prospective memory*, whereas the task-driven shifts of attention may have been amnesic, Horowitz et al. further concluded that some form of long-range memory is possible in visual search but that it comes at the cost of slow and inefficient shifts of attention. That was not the condition in our experiment, however. Our visual search task required shifts of attention as fast and as efficient as possible, in order to complete the search in the short amount of time permitted (4 sec). Moreover, shifts of attention were mainly task driven in our experiment, and yet they definitely carried some form of long-range memory.

SUPPLEMENTAL MATERIALS

The full text of Supplemental Appendix A, containing comparisons of our model with those of Aks et al. (2002), Peterson et al. (2001), and Najemnik and Geisler (2005), may be downloaded from <http://app.psychonomic-journals.org/content/supplemental>.

SUPPLEMENTAL APPENDIX A

In this Appendix A, we compare in particular our computational model and results with those of (1) an alternative approach to active visual search based on time series analysis and complex dynamics (Aks et al., 2002), (2) a previous set of models and simulations asserting that visual search has memory (Peterson et al., 2001), and (3) a contrasting perspective of an “ideal Bayesian observer” (Najemnik & Geisler, 2005). These comparisons further require reasonably detailed reviews and analyses of those three papers in particular and are thus provided as supplemental material accessible online at <http://app.psychonomic-journals.org/content/supplemental>. All references cited in Supplemental Appendix A are included among the present article’s references.

APPENDIX B
Local and Averaged SAD and ISD Distributions and Their Ratios

Here, we explain mathematically the meaning of probabilities associated with saccade amplitude and inter-stimulus distance distributions, their ratios, averaging, and local dependence.

Associated with the SAD, there is a probability $P_{\text{SAD}}(\mathbf{x}, r)dr$ of generating a saccade from the current fixation point at the tip of the two-dimensional vector \mathbf{x} to all other point vectors \mathbf{y} at distances between $r = |\mathbf{y} - \mathbf{x}|$ and $r + dr$, regardless of the relative angle between the \mathbf{x} and $\mathbf{y} - \mathbf{x}$ vectors. Likewise, associated with the ISD distribution, there is a probability $P_{\text{ISD}}(\mathbf{x}, r)dr$ of having another search item at any distance between $r = |\mathbf{y} - \mathbf{x}|$ and $r + dr$ from the original item at \mathbf{x} , regardless of the relative angle between the \mathbf{x} and $\mathbf{y} - \mathbf{x}$ vectors.

Now, suppose that search items were randomly and uniformly distributed throughout a *disklike display* of radius R . Then, $P_{\text{ISD}}(r)dr = 2\pi r dr / (\pi R^2)$ would clearly correspond to the ISD distribution, provided that the original item was located at the center of the disklike display. However, if the original item was located at any other point \mathbf{x} , then $P_{\text{ISD}}(\mathbf{x}, r)dr$ would have to change and depend on \mathbf{x} also, the more so the more \mathbf{x} moved away from the center of the disklike display. But we could still define an average $\langle P_{\text{ISD}}(r) \rangle dr$ that depends only on the r distance by averaging over all \mathbf{x} values for the possible location of the original item. That is exactly the procedure that is implemented computationally, in particular for the rectangular display in our experiment (see Figure 1 and its discussion in the present article). Now, for small r , the ISD distribution should still be proportional to r , as it would be in a boundless display, or in the disklike display with \mathbf{x} at its center, because the item distribution is still uniform and $2\pi r dr$ is just the area of the annulus of radius r and infinitesimal thickness dr . However, when r approaches and eventually exceeds the diagonal of our rectangular display, the ISD must evidently vanish, since there are no search items beyond that distance from any \mathbf{x} within the display, even if \mathbf{x} lies at one of the display corners. The initial proportionality to r , followed by the predicted decrease, are indeed evident in the computed ISD distribution shown in Figure 1.

Now, the $P_{\text{SAD}}(\mathbf{x}, r)/P_{\text{ISD}}(\mathbf{x}, r)$ ratio represents the probability of generating a saccade from the current fixation point at \mathbf{x} to just a single point \mathbf{y} at a distance $r = |\mathbf{y} - \mathbf{x}|$ away from \mathbf{x} . After averaging over \mathbf{x} , we obtain the probability factor $f(r) = \langle P_{\text{SAD}}(r) \rangle / \langle P_{\text{ISD}}(r) \rangle$ that enters the probability $p_{i \rightarrow j}$ of executing a saccade from a search item i to another specific item j , rather than to any other item that may lie at the same distance $r = |\mathbf{y} - \mathbf{x}|$ from the given i item. The selection of a single j item in the $p_{i \rightarrow j}$ probability is what prompts the introduction of the ISD distribution and the division by it to obtain the probability factor $f(r)$ —which is a ratio of distributions and not a distribution itself, by the way.

One could be further interested in studying the local dependence of $P_{\text{SAD}}(\mathbf{x}, r)$ and $P_{\text{ISD}}(\mathbf{x}, r)$ on the initial fixation location at \mathbf{x} . Motter and Holsapple (2007) have thoroughly examined that dependence by partitioning the rectangular display into various zones from which each saccade may originate (see their Figure 3A). Their Figures 3C and 3D then report the corresponding local ISDs and SADs. Remarkably, the $P_{\text{SAD}}(\mathbf{x}, r)/P_{\text{ISD}}(\mathbf{x}, r)$ ratio, which they define in their Equation 1 as the “probability of making a saccade by chance,” $P_S(r)$, times the number of search items, N , turns out to be rather independent of the initial fixation location at \mathbf{x} (see Figure 3E in Motter & Holsapple, 2007). Thus, even the $P_{\text{SAD}}(\mathbf{x}, r)/P_{\text{ISD}}(\mathbf{x}, r)$ ratio may essentially coincide with our averaged probability factor $f(r)$, independent of \mathbf{x} .

APPENDIX C

Relation Between the SAD and the Average Number of Saccades per Trial

Deeper understanding and further confirmation of the basic mechanisms underlying the conjunctive visual search that we have studied derive from an analytical formulation and a statistical estimate of the relation between two critical observations—namely, the SAD and the average number of saccades needed to capture the target. Here, we explain their deeper connection purely on the basis of the combined roles that the AC plays in attention and memory.

We start by considering the idealization of *perfect memory tagging* and assume a corresponding model of *sampling without replacement*. We thus generate an initial estimate of the number of fixations F_I required to capture the target by simply dividing the number N of relevant objects (having the same color as the target) in the whole display by the average number of relevant objects that are processed in parallel during each fixation. The latter number equals the AC times the density ρ of relevant objects. The radius R_{AC} of the AC may be further expressed as the c-ANND times the extension coefficient, determined from computer simulations and shown in Table 2. An extension coefficient smaller than 1 (for 12-distractor arrays) implies that only one relevant object is sampled during each fixation. In the present context and equations, such an extension coefficient is equated to 1 numerically. The estimated number of fixations required to inspect all the N relevant objects is thus given by

$$F_I = \frac{N}{\rho \cdot \pi \cdot R_{AC}^2} = \frac{\text{Screen Area}}{\text{AC}}. \quad (\text{C1})$$

Namely, F_I represents the number of samples of area $\text{AC} = \pi \cdot R_{AC}^2$ needed to cover the entire display area, which is $16.3^\circ \times 24.4^\circ$. Sampling without replacement, the average over many trials of the number of fixations required to capture the target is about one half the number F_I required to inspect all relevant objects. The corresponding average number of saccades required to capture the target then is

$$S_C = \frac{1}{2} F_I - 1. \quad (\text{C2})$$

Figure C1 shows the prediction by this initial estimate for each array size searched by young (panel A) and older (panel B) adults. Evidently, this prediction underestimates systematically the measured values. The reason is that the number of fixations F_I estimated in Equation C1 requires sampling of a different set of relevant objects during each fixation. In turn, that requires that the distance between consecutive fixations—hence, the amplitude of the intervening saccade—be at least twice as large as the AC radius. However, it is clear from Figure 2 and others showing measured saccade amplitude distributions that many saccades have amplitudes less than twice the AC radius. This implies that some relevant objects within an AC may be sampled again during the following fixation. As a result, more fixations are typically required to sample the entire set of relevant objects than F_I estimated in Equation C1.

Removing the number of relevant objects that are sampled once more during consecutive fixations requires estimating the overlap of consecutive ACs during the search. Let us then consider two circles, denoted as F_1 and F_2 , with the same radius r and centers separated by a distance d , as shown in Figure C2. The area of F_2 outside of F_1 equals the area of F_2 minus the “lens area” $F_1 \cap F_2$, which is in turn given by

$$\text{AL} = 2r^2 \text{Cos}^{-1}(d/2r) - \frac{d}{2} \sqrt{(-d+2r)(d+2r)}. \quad (\text{C3})$$

By varying d_i between 1.4° and 25° , in increments of 0.1° , in the abscissas of saccade amplitude distributions, such as those shown in Figure 2, we can determine the fraction SA_i of saccades whose amplitudes have d_i values within those increments. Using Equation C3, we can then find the corresponding lens areas AL_i . So, we can replace the AC in Equation C1 with the more appropriate SA_i -weighted average of $(\text{AC} - \text{AL}_i)$, yielding

$$F_I = \frac{N}{\rho \cdot \sum_{i=14}^{250} (\text{AC} - \text{AL}_i) \cdot SA_i}. \quad (\text{C4})$$

With this refined estimate of F_I , Equation C2 should provide a more reasonable prediction of the average number of saccades S_C per trial.

Figures C3A and C3B show the measured average number of saccades needed to capture the target and the corresponding prediction based on Equation C4 for each array size in both age groups. Their reasonably good agreement confirms that our basic understanding of the roles played by the AC in both attentional sampling and the corresponding memory tagging of relevant objects allows us to predict reasonably well the observed average number of saccades per trial from the observed saccade amplitude distribution, for all array sizes in both age groups.

APPENDIX C (Continued)

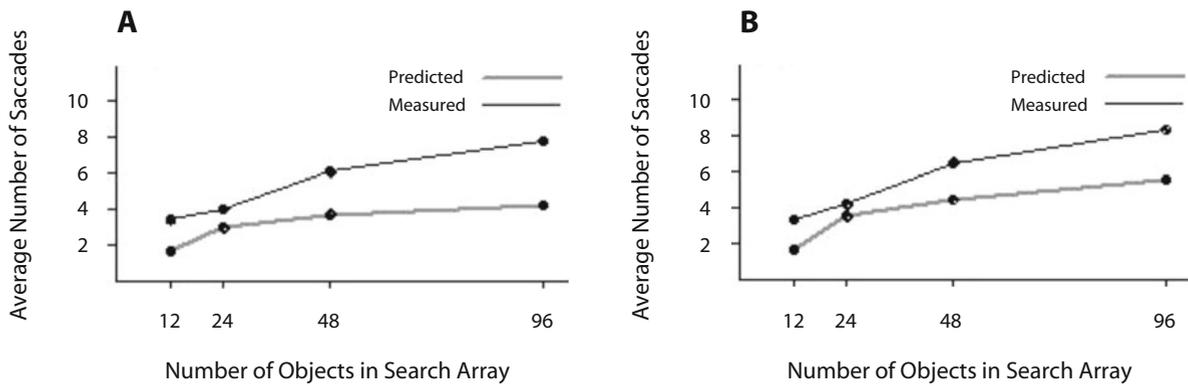


Figure C1. Measured and predicted numbers of saccades needed to capture the target, assuming sampling without replacement, for young (A) and older (B) adults.

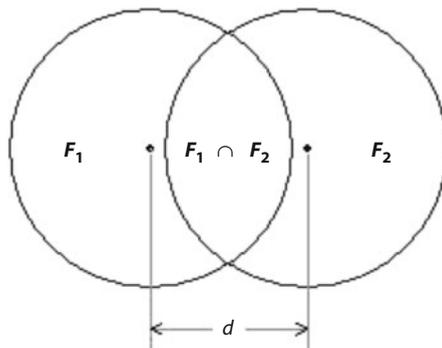


Figure C2. Area of circle-circle intersection.

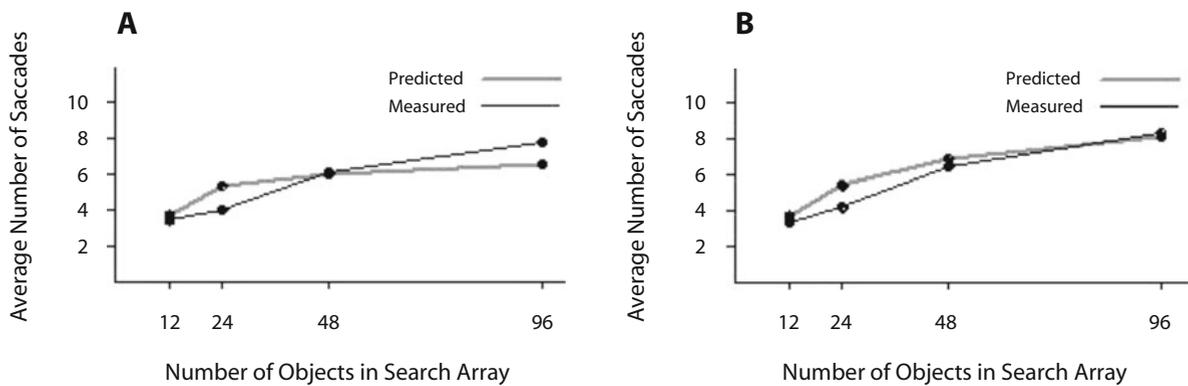


Figure C3. Measured and predicted average numbers of saccades needed to capture the target, assuming sampling with area lens replacement, for young (A) and older (B) adults.