Eye-Movement Analysis of Search Effectiveness

Ralf VAN DER LANS, Rik PIETERS, and Michel WEDEL

Advances in eye-tracking technology have promoted its widespread use to understand and improve target searches in psychology, industrial engineering, human factors, medical diagnostics, and marketing. Eye movements are the realization of a complex, unobserved spatiotemporal attention process with many sources of variation. Eye-tracking data often have been aggregated and/or summarized descriptively, because few adequate statistical models are available for their analysis. This article proposes a model that may serve to uncover the latent attention processes of people searching for targets in complex scenes. It recognizes the spatial nature of eye movements and represents two latent attention states, a localization state and an identification state, between which people may switch over time according to a Markov process. A saliency map, based on low-level perceptual features and the scene's organization, guide target searches in the localization state. In the identification state, people verify whether a selected candidate object is the target. The model is applied to analyze commercial eye-tracking data from more than 100 people engaged in a target search task on a computer-simulated retail shelf display. Rapid switching between attention states over time is revealed. Estimates of the feature and saliency maps are provided and found to be related to search performance. The results facilitate the evaluation of the effectiveness of alternative visual search strategies.

KEY WORDS: Gibbs sampling; Hidden Markov model; Hierarchical model; Slice sampling; Spatial analysis.

1. INTRODUCTION

Target search is one of the most common and important tasks that people perform daily; for example, radiologists search for faint nodules in chest radiographs, air traffic controllers search for aircraft symbols that pose a potential collision threat on their radar screens, airport security personnel search for concealed weapons in X-ray images of luggage, car drivers search for traffic signs, and consumers search for products on websites and overstocked shelves of retail outlets. Because search tasks prevail in daily life, inaccuracies and slow response can have serious implications. There is a long tradition of research on target searches in psychology and other fields such as industrial engineering, human factors, and medical diagnostics (Rayner 1998; Wolfe 1998; Ho, Scialfa, Caird, and Graw 2001; Vora et al. 2002; McCarley, Kramer, Wickens, Vidoni, and Boot 2004). Such research aims to improve the selection and training of human search agents and the content and organization of search displays, such as shelves and websites (Gramopadhye, Drury, and Sharit 1997; Wang, Lin, and Drury 1997; Yang, Dempere-Marco, Hu, and Rowe 2002). A thorough understanding of target search is required for those purposes. The goal of the present study is to develop a statistical model, calibrated on eye-tracking data, to contribute to that understanding and evaluate the effectiveness of visual search strategies.

The traditional approach has been to infer what people pay attention to during searches from response accuracies and latencies in search tasks (Pashler, Johnston, and Ruthruff 2001). This has proven difficult, because attention is an unobservable mental process, different attentional processes may lead to the same search performance, and the same process may produce different performances due to individual differences (Sanders and Donk 1996; Pashler 1998). Therefore, research has focused on basic search tasks for abstract stimuli in simple multielement displays, rather than on searches for realistic targets in the complex scenes that people encounter daily. This has hampered progress in theory development and applications, because it is not obvious that the findings thus obtained generalize to such natural situations (Bülthoff and Veen 2001; Kingstone, Smilek, Ristic, Friesen, and Eastwood 2003). This has led to a recent surge in scientific interest in studying visual attention to complex scenes and in the application of eye-movement recordings to investigate target search (Yang et al. 2002; Duchowski 2003).

Eye movements constitute measures of the unobserved visual attention process with a high temporal and spatial resolution, and thus have the potential to yield insights about target search that are hard to obtain otherwise (Findlay and Gilchrist 1998). Current eye-tracking technology enables examination of comparatively large samples of individuals (Pieters and Wedel 2004), which facilitates quantitative analysis. But much of the previous work has relied either on descriptive statistics of eye movements or used models of only selected aspects of target search (Zelinsky 1996; Inhoff and Radach 1998; Rayner 1998; Motter and Holsapple 2001; Duchowski 2003). Despite their potential to provide better understanding, comprehensive statistical models of the spatiotemporal attentional processes underlying target search have not been developed. This may be due at least in part to the complexity and computational requirements of the modeling task. Here we propose and test such a model for eye-movement analysis of search effectiveness.

2. EYE MOVEMENTS DURING TARGET SEARCH

Visual search is the process of locating a target among a set of distractors in a scene (Wolfe 1998). The target can be any object, like a red square, a bone fracture, a dot on a radar screen, a splinter in a lens, or a product on a shelf. The distractors are all other objects in the scene that are, not the target. In real life searches, targets often are neither uniquely defined nor distinguished from distractors by only a few perceptual features, and the distractors usually vary considerably among themselves as well. These factors complicate search (Duncan and Humphreys 1989), and attention is required to reduce the uncertainty about

Ralf van der Lans is Assistant Professor, Department of Marketing Management, Rotterdam School of Management Erasmus University, 3000 DR Rotterdam, The Netherlands (E-mail: *RLans@rsm.nl*). Rik Pieters is Professor, Marketing Department, School of Economics and Business, Tilburg University, 5000 LE Tilburg, The Netherlands (E-mail: *Pieters@uvt.nl*). Michel Wedel is Pepsico Professor of Consumer Science, Department of Marketing, Smith School of Business, University of Maryland, College Park, MD 20742 (E-mail: *mwedel@rhsmith.umd.edu*). The authors thank the editor and two anonymous referees for their helpful suggestions on a previous version of the article. They also thank Dennis Fok, Harald van Heerde, and Richard Paap for valuable suggestions. Verify International is gratefully acknowledged for the eye-tracking database and their support during the project.

Journal of the American Statistical Association
June 2008, Vol. 103, No. 482, Applications and Case Studies
DOI 10.1198/016214507000000437

both the location and the identity of the target. Location uncertainty pertains to where in the scene the target is located, and identity uncertainty pertains to whether an object that has been located is the search target or a distractor.

Visual search thus is an active process that involves attention processes, which invoke movements of the eyes across the scene to reduce location and identity uncertainties over time. Eye movements on stationary scenes consist primarily of fixations and saccades (Rayner 1998). Fixations are brief moments in which the eye is relatively stable, and serve to project an object or region in the scene through the line of sight onto the fovea, the small area of the retina with the highest acuity. During an eye fixation, information is extracted from the perceptual field around the exact fixation position (Anstis 1974; Sanders and Donk 1996). Saccades are rapid ballistic movements of the eyes—during which vision is suppressed—to redirect the line of sight to a new fixation position. The statistical challenge is to identify the unobservable spatiotemporal attention processes during target search using information on the pattern of eye fixations. In our statistical model, the sequence of eye-fixation positions are the units of analysis, because information is extracted at the fixation locations.

2.1 Attention Switching During Target Search

Because human information-processing capacity is limited, efficient attentional mechanisms need to select the most relevant information from the scene at any point in time. This involves two latent attention states in which the reduction of location uncertainty ("where") or identity uncertainty ("what") prevails (Niebur and Koch 1998). There is ample evidence for separate neural pathways in the human brain for object localization and object identification. These are the dorsal (or "where") stream, passing from the primary visual cortex into the parietal lobe, and the ventral (or "what") stream, terminating in the temporal lobe (Ungerleider and Mishkin 1982). Neuropsychological research indicates that these separate pathways may produce different eye-movement patterns, with more dense fixations during identification (Bullier, Schall, and Morel 1996; Thompson 2005).

2.1.1 Localization State. The visual brain decomposes visual information from the image projected on the retina into separate maps, representing the basic perceptual features: color, luminance, and edges. These feature maps, arising bottom-up from the scene, are processed in specialized areas of the primary visual cortex that exhibit a detailed topographic representation of the visual field (Hubel and Wiesel 1962; Zeki 1993; Fuster 2003). The visual brain builds a saliency map of the scene as a weighted combination of the feature maps, with the weights arising top-down from the search task and knowledge of the person searching (Treisman and Gelade 1980; Wolfe 1998). The saliency map is thought to be maintained in specialized brain structures also involved in the motor control of eye movements. In the localization state, the saliency map guides the focus of attention (FOA) to quickly select scene regions with candidate targets. Winner-takes-all and inhibition-of-return mechanisms deploy attention to visit locations in order of their saliency (Itti and Koch 2001; Pomplun, Reingold, and Shen 2003).

Attention also may be guided systematically by the scene's organization. On exposure, the search display is rapidly segmented into its constituent objects and their spatial arrangement (Duncan and Humphreys 1989; Wertheimer 2001). Attention is then expressed in systematic, orderly sequences of fixations and saccades across those scene segments (objects and regions) (Ponsoda, Scott, and Findlay 1995; Horowitz and Wolfe 2003). For instance, Monk (1984) observed systematic horizontal zigzag patterns (left–right and right–left) in eye movements during target search across a regular multielement display.

These saliency and systematic search strategies (Horowitz and Wolfe 2003) guide attention to localize candidate targets in the scene. Little is known, however, about the relative importance of different perceptual features in the formation of the saliency map, the prominence of systematic search, and the effectiveness of these search strategies in complex scenes.

2.1.2 Identification State. Once a candidate is localized, attention switches to the identification state. Then the candidate object is matched to a template of the target in memory from which detailed information must be sampled. Because complex objects in cluttered scenes cannot be identified in a single fixation, refixation is needed to verify their identity (Henderson, Weeks, and Hollingworth 1999). Search terminates in the identification state when sufficient evidence is available for a positive match, and continues in the localization state in the event of a negative match. No comprehensive statistical models have been developed that allow these attention processes to be captured from eye-movement data.

3. DATA DESCRIPTION

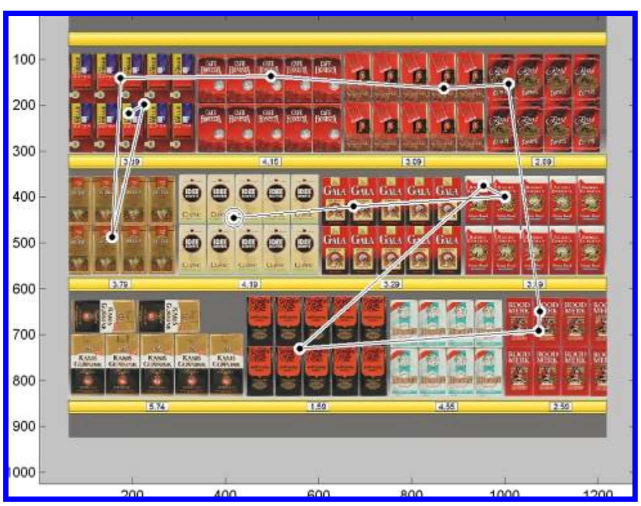
We analyze eye-movement data collected by a marketing research company in a brand search experiment. Consumers engage in brand searches almost every day, yet brand search turns out to be a challenging task that merits more investigation (Drèze, Hoch, and Purk 1994).

3.1 Experiment

A total of 106 randomly sampled consumers (52% female and 48% male; age range 16–55 years) participated in the experiment. Participants were instructed to search for a specific brand (providing the brand name as a cue) of coffee among a set of 12 different existing coffee brands on a computer-simulated supermarket shelf. They indicated that they found the target by touching it on the touch-sensitive screen. The instructions and search display were presented on NEC 21-inch LCD monitors, and participants were seated in front of the screen. The search display was shown full-screen at a resolution of 1,024 × 1,280 pixels (approximately $15.5^\circ \times 19.4^\circ$ of visual angle) in full-color mode. It contained 12 brand groups, with size ranging between $3.0^\circ \times 3.6^\circ$ and $3.0^\circ \times 5.8^\circ$ of visual angle, and multiple replications of each brand ("facings") in these groups (109 in total). The display was presented for a maximum of 10 seconds, which is realistic for in-store product and brand search (Hoyer 1984). During the search task, participants' eye movements were recorded with infrared corneal reflection eye-tracking methodology, with a temporal resolution of 20 ms and spatial resolution better than 0.5° (Wedel and Pieters 2000; Duchowski 2003). The specific eye-tracking methodology allows participants to freely move their head within a virtual box of about 20 inches, while cameras track the eyes and head continuously. For all 106 participants, the complete pattern of fixations and saccades across time and the search display was retained.

3.2 Data Structure

Figure 1 presents an example of the eye-movement pattern (top) and the corresponding data for a specific individual (bottom). The brands share perceptual features, such as shape and colors, and there are multiple packages in each brand group, as shown in Figure 1, making the search task difficult. The eye-movement data consist of the coordinates of the sequence of fixations on the LCD display. This makes it possible to relate the fixation positions to image characteristics at the pixel level, such as its RGB color values, the brand group to which it corresponds, and whether or not it represents a text element, such as the brand name. We define the data in terms of these characteristics of the exact fixation positions. For example, in Figure 1, the first fixation is on pixel (445, 418), which belongs to brand group 6, is not a text element, and has RGB values (189, 157, 106).



Fixation	Row (r)	Column (c)	Brand group	Text	R	G	B
1	445	418	6	0	189	157	106
2	421	675	7	0	113	112	110
3	400	1,000	8	1	192	192	192
...	...	...	...	...	...	...	...
13	199	225	1	0	51	31	59
14	217	192	1	1	145	56	48

Figure 1. Computer-simulated shelf with an observed pattern of eye movements. The circle in the top figure indicates the first eye fixation. The target is in the top left of the display. An observed pattern of eye movements of a single individual is superimposed on the display. It consists of 14 eye fixations (dots), connected with 13 saccades (lines connecting the dots). The table indicates for each eye fixation its exact pixel position, as the display is represented as a matrix containing 1,024 × 1,280 pixels. For each pixel, we code to the brand group to which it belongs, whether it is part of the text (1 when pixel is part of the text, 0 otherwise), and its RGB color value.

3.3 Decomposition of the Search Display

We decompose the search display (Fig. 1) in basic features and segments/objects to determine the role of the saliency and systematic strategies in target search. We use the basic perceptual features, color, luminance, and edges (Wolfe 1998; Itti and Koch 2001), and coded as segments of the image the shelf, the 12 brand groups, and text elements in the brand groups. For each pixel in the display, the RGB values and the object and text to which it belongs are known, coded as $1{,}024 \times 1{,}280$ matrices. The RGB values in each pixel are used to define color, luminance, and edges. We code color features for red, gold and blue, because these colors differ systematically in diagnosticity in the current search display and task. Red has low diagnosticity because many brands share it; gold and blue have high diagnosticity because few objects share it, with gold having negative (absent in target) and blue having positive diagnosticity (present in target). Colors are coded as dummy variables at the level of pixels. Luminance is computed as the weighted sum of the three RGB values, following the NTSC and JPEG standards (Gevers 2001): $S_{\text{luminance}} = .299R + .587G + .114B$. The luminance values are then used to compute the edges of objects, because edges are sharp changes in luminance. In the visual brain, these edges assist in scene segmentation. Following Marr (1982), we define edges at locations with a maximum change in luminance, which occurs where the second-order derivative of luminance equals 0, and retain those corresponding to the borders of the shelf, brand groups, and text elements in the display. Shelf, brand, and text segments are coded as dummy variables at the level of pixels, that is, as 1 if the corresponding pixel belongs to the specific segment, and 0 otherwise.

4. THE MODEL

Let $S = \{S_1, S_2, \dots, S_M\}$ represent M basic features (color, luminance, edges) of the search display D. Each S_m, with $m \in 1, \dots, M$, represents a $R \times C$ matrix that contains the value of feature m for each pixel (in the application, $R = 1{,}024$ and $C = 1{,}280$; see Sec. 3.3). Similarly, $\mathbf{O} = \{O_1, O_2, \dots, O_Q\}$ represents Q different $R \times C$ matrices representing image segments (shelf, brand groups, and text elements) of the display, where $O_q(r, c) = 1$ if pixel coordinate (r, c) belongs to segment q (with $q \in 1, \dots, Q$). Because visual acuity declines progressively with increasing eccentricity from the current ith fixation position (r_{hi}, c_{hi}) for individual h (Anstis 1974)—that is, from foveal to peripheral vision—we assume that the perceptual field from which useful information is extracted is represented by a bivariate normal distribution around (r_{hi}, c_{hi}) with a standard deviation equal to ζ. Thus the value of $S_m(r, c, \zeta)$ in location (r, c) with perceptual field ζ is defined as

$$S_m(r, c, \zeta) = \frac{1}{\sqrt{2\pi}\,\zeta} \int_{(a,b)\in D} S_m(a, b, 0) \times \exp\left(-\frac{(a-r)^2 + (b-c)^2}{2\zeta^2}\right) \mathrm{d}a\,\mathrm{d}b. \quad (1)$$

We define $O_q(r, c, \zeta)$ in a similar way. [We suppress ζ in $S_m(r, c, \zeta)$ and $O_q(r, c, \zeta)$ for simplicity of notation, where possible.] The perceptual field acts as a spatial smoother with a Normal kernel and bandwidth of ζ. Previous research has shown that the perceptual field is between 1 (Pomplun, Reingold, Shen, and Williams 2000; Motter and Holsapple 2001) and 2 degrees of visual angle (Rayner 1998), which corresponds approximately to foveal vision. Setting ζ to these values, however, accommodates peripheral vision with lower levels of visual acuity as well. In the empirical application, we compare models with several prespecified values of ζ.

If the ith fixation of individual h is at pixel location (r_{hi}, c_{hi}), then its corresponding value for feature m equals $S_m(r_{hi}, c_{hi})$ and that for image segment q equals $O_q(r_{hi}, c_{hi})$. In the model we use $O_{w_t(\tilde{q}(r_{h,i-1}, c_{h,i-1}))}(r_{hi}, c_{hi})$, where $w_t(q)$, $t = 1, \dots, T$ maps each image segment to the subsequent segment as dictated by the tth predefined search strategies and where $\tilde{q}(r, c) = q \in 1, \dots, Q$ indicates to which segment each pixel belongs. Refixations on the same image segment are considered special cases in which $w_t(q) = q$. We construct, for each i, h and (r, c), a $K \times 1$ pixel information vector $v_{hi}(r, c)$, where $K = 1 + M + T$, defined as

$$v_{hi}(r, c) = \Big[1, \overbrace{S_1(r, c), \dots, S_M(r, c)}^{M \text{ Basic Features}}, \overbrace{O_{w_1(\tilde{q}(r_{h,i-1}, c_{h,i-1}))}(r, c), \dots, O_{w_T(\tilde{q}(r_{h,i-1}, c_{h,i-1}))}(r, c)}^{T \text{ Image Segments}}\Big]. \tag{2}$$

The sequence $X_h = (r_{h1}, c_{h1}), \dots, (r_{h,n_h}, c_{h,n_h})$, with n_h indicating the last eye fixation of individual h, is assumed to arise from a spatial point process (Cressie 1993) on the display D, with intensity functions $\lambda_j(\theta_{jh}, v_{hi}(r, c))$ for each of two attention states j, $j = \{1, 2\}$, representing the localization and identification attention states. (We suppress the subsript i on j_i for simplicity of notation, where possible.) Note that these are *unobserved* states of the attention process. The dynamics of the transition between these two states are governed by a Markov process, which probabilistically affects the observed pattern of eye movements. Top-down processes influence the construction of the saliency map and the systematic search strategies in the localization state, through weights on the feature and segment inputs from the search display, $v_{hi}(r, c)$, represented by the vector $\theta_{jh} = \{\theta_{jh1}, \theta_{jh2}, \dots, \theta_{jhK}\}$. In the localization state (i.e., $j = 1$), all K parameters are estimated, because both saliency and systematic search may operate. In the identification state, only the parameters pertaining to refixations on the same image segments (text elements and brand groups) are estimated $(\theta_{2hK-1}, \theta_{2hK})$; all other parameters are set to 0.

Because eye-movement patterns vary substantially between individuals (Rayner 1998), we assume a normal distribution for the individual parameters, $\theta_{jh} \sim \mathrm{N}(\mu_j, \Sigma_j)$, with Σ_j diagonal. We assume that individuals switch between the two latent attention states according to a Markov process with transition matrix Π (Liechty, Pieters, and Wedel 2003). For each observed fixation point (r_{hi}, c_{hi}), the intensity function over the search display $\int_{(a,b)\in D} \lambda_j(\theta_{jh}, v_{hi}(a, b))\, \mathrm{d}a\, \mathrm{d}b$ integrates to 1. Thus $\lambda_j(\theta_{jh}, v_{hi}(r, c))$ can be interpreted as the probability that fixation i of individual h will be directed to pixel (r, c), given that it is generated in attention state j.

We use the following formulation for the intensity function:

$$\lambda_j(\theta_{jh}, v_{hi}(r, c)) = \frac{(v_{hi}(r, c) \cdot \theta_{jh})^2}{\int_{(a,b)\in D} (v_{hi}(a, b) \cdot \theta_{jh})^2\, \mathrm{d}a\, \mathrm{d}b} = \frac{(v_{hi}(r, c) \cdot \theta_{jh})^2}{\theta'_{jh} V_{hi} \theta_{jh}}, \tag{3}$$

with V_{hi} a $K \times K$ matrix in which element (k_1, k_2) represents the cross-product between pixel information elements k_1 and k_2 over the display D, that is,

$$V_{hi}[k_1, k_2] = \int_{(a,b)\in D} \big(v_{hi,k_1}(a, b) \cdot v_{hi,k_2}(a, b)\big)\, \mathrm{d}a\, \mathrm{d}b. \tag{4}$$

The square-root link function in (3) ensures nonnegativity of $\lambda_j(\cdot)$ and is theoretically appealing because it formulates the spatial intensity on the surface of the display $\lambda_j(\cdot)$ as the square of the (weighted) sum of the one-dimensional features and image segments contained in the pixel information vector $v_{hi}(r, c)$. In addition, it renders the model computationally feasible, because it simplifies the computation of the normalizing constant in (3) because, as shown in (4), V_{hi} does not depend on the parameter values.

Collecting all parameters in Θ, given the observed sequence $X \equiv (X_1, \dots, X_H)$, with H representing the total number of participants in the experiment, we have the following hidden Markov model (HMM) likelihood (Scott 2002):

$$L(X; \Theta) = \prod_{h=1}^{H} \left(\left\{ \sum_{j_2=1}^{2} \cdots \sum_{j_{n_h}=1}^{2} \prod_{i=2}^{n_h} \pi_{j_{i-1}, j_i} \times \lambda_{j_i}(\theta_{jh}, v_{hi}(r_{hi}, c_{hi})) \right\} \prod_{j=1}^{2} \phi(\theta_{jh}, \mu_j, \Sigma_j) \right), \tag{5}$$

satisfying

$$\int_{(a,b)\in D} \lambda_{j_i}(\theta_{jh}, v_{hi}(a, b))\, \mathrm{d}a\, \mathrm{d}b = 1 \quad \forall j_i = 1, 2, \forall h = 1, \dots, H, \forall i = 2, \dots, n_h. \tag{6}$$

We assume, to identify the model, that the first fixation of each individual is made in the localization state (i.e., $j_1 = 1$), because it is thought to be used by the visual brain to rapidly extract features and segments from the image (Friedman 1979; Henderson et al. 1999). Therefore, $\mathbf{S}$ and $\mathbf{O}$ do not guide the first fixation, and its position is not taken into account in estimating θ_{jh}. Restriction (6), which is satisfied by (3), induces the loss of 1 degree of freedom, so that we constrain the constant θ_{jh1}, for each j and h to be equal to 1.

4.1 Estimation

The model is estimated using Markov chain Monte Carlo (MCMC) methods. Combining (3)–(6) with conjugate prior distributions for μ_j [i.e., $\mu_j \sim \mathrm{N}(\eta_j, L_j)$] and Σ_j [i.e. $\Sigma_j^{-1} \sim$

$W(G_j, g_j)$, $j = 1, 2]$ leads to the following conditional posterior distribution for Θ:

$$p(\Theta|X) \propto \prod_{h=1}^{H} \prod_{j=1}^{2} \prod_{i:z_{hi}=j}^{n_h} \lambda_j\big(\theta_{jh}, v_{hi}(r_{hi}, c_{hi})\big) \times p(\theta_{jh}|\mu_j, \Sigma_j^{-1}) \cdot p(\mu_j) \cdot p(\Sigma_j^{-1}). \quad (7)$$

In (7), the latent variables $z_{hi} \in \{1, 2\}$ indicate the attention state from which fixation i of individual h is generated (Robert, Celeux, and Diebolt 1993); the z_{hi}'s on which (7) is conditioned constitute a Markov chain with transition matrix $\Pi = (\pi_{j_1,j_2})$, with $\pi_{j_1,j_2} = P(z_{hi} = j_2|z_{h,i-1} = j_1)$.

4.1.1 Conditional Posterior for θ. Using (7), we write the posterior for θ_{jhk} as

$$p(\theta_{jhk}|\theta_{jh,-k}, \mu_j, \Sigma_j^{-1}, X) \propto \phi(\theta_{jhk}, \mu_{jk}, \sigma_{jkk}) \cdot \prod_{i:z_{hi}=j}^{n_h} \lambda_j\big(\theta_{jh}, v_{hi}(r_{hi}, c_{hi})\big). \quad (8)$$

Because $\lambda_j(\theta_{jh}, v_{hi}(r_{hi}, c_{hi})) > 0$ for every eye fixation i, we apply augmented variable Gibbs sampling (Damien, Wakefield, and Walker 1999; Neal 2003). Toward this end, we introduce auxiliary variables u_{hi} drawn uniformly from the interval $[0, \lambda_j(\theta_{jh}, v_{hi}(r_{hi}, c_{hi}))]$. Conditional on these, the individual parameters θ_{jhk}, $k = 2, \ldots, K$, are drawn sequentially from truncated normal distributions (Robert 1995) such that

$$u_{hi} < \lambda_j\big(\theta_{jh}, v_{hi}(r_{hi}, c_{hi})\big) \qquad \forall i, z_{hi} = j. \quad (9)$$

The allowed region implied by (9) is determined analytically.

4.1.2 Conditional Posterior for μ and Σ^{-1}. The conditional posteriors for μ_j and Σ_j^{-1} are standard,

$$\mu_j|\cdots \sim \mathrm{N}\left(B_j\left(\Sigma_j^{-1}\sum_{h=1}^{H}\theta_{jh} + L_j^{-1}\eta_j\right), B_j\right) \quad (10)$$

and

$$\Sigma_j^{-1}|\cdots \sim \mathrm{W}\left(\left(\sum_{h=1}^{H}(\theta_{jh} - \mu_j)(\theta_{jh} - \mu_j)' + G_j^{-1}\right), H + g_j\right), \quad (11)$$

where $B_j = (H\Sigma_j^{-1} + L_j^{-1})^{-1}$.

4.1.3 Estimation of z and Π. We draw the missing variables z_{hi} from a hidden Markov chain with transition probabilities $\pi_{j_1,j_2} = P(z_{hi} = j_2|z_{h,i-1} = j_1)$. We postulate a Dirichlet prior distribution $\Pi_{j\cdot} \sim D(\Xi_{j\cdot})$ for each row j of Π, with $\Xi = \binom{\xi_{11}\ \xi_{12}}{\xi_{21}\ \xi_{22}}$ and use the forward–backward recursion algorithm to estimate z (Chib 1996; Scott 2002). Given z, the posterior distribution of $\Pi_{j\cdot}$ is given by

$$\Pi_{j\cdot}|\cdots \sim \mathrm{D}\Bigg(\xi_{j1} + \sum_{h=1}^{H}\sum_{i=2}^{n_h} I\{z_{h,i-1} = j\} \cdot I\{z_{hi} = 1\}, \; \xi_{j2} + \sum_{h=1}^{H}\sum_{i=2}^{n_h} I\{z_{h,i-1} = j\} \cdot I\{z_{hi} = 2\}\Bigg), \quad (12)$$

with $I\{\cdot\}$ the indicator function.

The MCMC is run with a burn-in of 10,000 and 2,500 target draws thinned 1 in 2, convergence is monitored using standard methods. We standardize the data such that

$$\int_{(a,b)\in D} (v_{hik}(a, b))^2 \,\mathrm{d}a\,\mathrm{d}b = 1 \qquad \forall k = 1, \ldots, K,$$

and

$$\int_{(a,b)\in D} v_{hik}(a, b) \cdot v_{hi1}(a, b) \,\mathrm{d}a\,\mathrm{d}b = 0 \qquad \forall k = 2, \ldots, K.$$

In the identification state, we restrict all parameters to 0, except for those corresponding to the refixations on the image segments. Thus, label switching does not occur, because the two hidden Markov states are differently parameterized (Frühwirth-Schnatter 2001). We use diffuse priors, $\eta_j = 0$, $L_j = \mathrm{diag}(1{,}000)$, $G_j = \mathrm{diag}(1)$, $g_j = K_j + 1$, and $\Xi = \binom{1\ 1}{1\ 1}$, with K_j the number of free parameters in attention state j. Analysis of synthetic data, reported in Appendix A, reveal that the chains converge before the burn-in and recover the true parameters well within twice the posterior standard deviation. We compare models with different perceptual fields by computing the log-marginal densities of the data with the method of Chib (1995).

5. RESULTS

5.1 Saliency and Systematic Search

We investigated the size of the perceptual field by estimating models for a range of values of ζ, and comparing the log-marginal densities (LMDs) based on the work of Chib (1995). We obtained the following LMD values for different perceptual fields: $\zeta = 1.0$, LMD $= -25{,}525$; $\zeta = 1.6$, LMD $= -25{,}479$; $\zeta = 1.8$, LMD $= -25{,}464$; $\zeta = 2.0$, LMD $= -25{,}497$; and $\zeta = 2.3$, LMD $= -25{,}576$. The maximum occurs at $\zeta = 1.8°$ of visual angle, for which we present the results next.

Table 1 presents the parameter estimates of the feature and saliency maps, as well as those representing systematic search, in the localization state. The color red is the "category code" shared by most brands in this category, including the target, and thus is not very diagnostic. The credible interval of its parameter does not cover 0, with a median estimate of .25, indicating that this feature strongly contributes to the salience map and guides attention in the localization state. The negatively diagnostic color gold (absent in the target) appears to attract attention as well, although less strongly, as indicated by a posterior median estimate of its parameter of .18. All draws from the posterior distribution of this parameter are positive. Interestingly, the posterior median of the positively diagnostic color blue (present in target) equals .38, and 97% of the posterior draws are larger than those of the category code red. This reveals the strong top-down weight placed on this feature during search. The negative posterior median for luminance (−.16) shows that the FOA is directed to the darker regions and away from the brighter regions of the display. The posterior medians of the standard deviations of the heterogeneity distributions are relatively large. This indicates that individuals differ substantially in the weights they place on the perceptual features top-down, probably because they retrieve from memory different templates of the target. The posterior median of the

Table 1. Parameter estimates of overall means and variances for search strategies: Median and credible intervals

	μ percentiles			σ percentiles		
Parameters	.025	.500	.975	.025	.500	.975
Localization state						
Refixation brand	.08	.26	.40	.17	.23	.31
Refixation text	−.42	−.25	−.13	.18	.23	.30
Saliency						
Color						
Red	.12	.25	.36	.15	.19	.24
Gold	.07	.18	.29	.15	.18	.22
Blue	.31	.38	.45	.16	.20	.25
Luminance	−.25	−.16	−.08	.15	.19	.25
Edges						
Brand group	−.38	−.28	−.19	.17	.23	.29
Shelf	.35	.51	.69	.16	.20	.25
Systematic						
Horizontal zigzag						
Left–right	.39	.47	.56	.19	.25	.32
Right–left	.43	.51	.59	.16	.20	.25
Identification state						
Refixation brand	2.63	3.33	4.07	.27	.43	.69
Refixation text	1.83	2.55	3.43	.29	.45	.74

parameter for shelf edges is positive (.51), indicating that people segment the scene based on the shelf layout to help guide attention (Wertheimer 2001). Quite interestingly, the posterior median of the edge parameters corresponding to brand groups is negative (−.28). This reveals that individuals actively direct their FOA away from the edges and to the center of the brand groups, where diagnostic information is expected to reside.

Using the posterior medians of the parameters for the perceptual features, we can now estimate a saliency map for each individual separately. Figure 2 presents the aggregate saliency map and the feature maps on which it is based. Table 2 presents the relative saliency of each brand group on the shelf, which is computed as

$$\text{relative saliency object } g = \frac{1}{W}\sum_{w=1}^{W} \frac{\int_{(a,b)\in g} \lambda_1(\mu_{1h}^{(w)}, v_{hi}(a,b))\,\mathrm{d}a\,\mathrm{d}b}{\sum_{q\in\tilde{Q}} \int_{(a,b)\in q} \lambda_1(\mu_{1h}^{(w)}, v_{hi}(a,b))\,\mathrm{d}a\,\mathrm{d}b}, \quad (13)$$

Table 2. Relative saliency of the 12 brands on the shelf

	Percentiles*				Percentiles		
Brand	.025	.500	.975	Brand	.025	.500	.975
1	2.89	3.37	3.89	7	.78	.86	.94
2	1.09	1.20	1.33	8	1.06	1.20	1.37
3	.97	1.08	1.21	9	.31	.40	.53
4	.70	.81	.92	10	.72	.88	1.04
5	.51	.71	.92	11	.14	.30	.50
6	.25	.41	.59	12	.70	.81	.93
Shelf	1.00	1.00	1.00	Borders	.08	.10	.13

* Values >1 indicate relatively high saliency; values <1 indicate relatively low saliency.

where $w = 1, \ldots, W$ represent the target draws from the MCMC sampler and $\tilde{Q}$ represents the subset of Q representing the brand groups. In Table 2 values >1 correspond to brands with a relatively high saliency. Whereas the relative saliency of the target brand ($g = 1$) is high (3.37), a number of distractor brands ($g = 4, 5, 6, 7, 9, 11, 12$) have saliencies significantly lower than 1.00. Thus the saliency measure (13) effectively assesses the perceptual "pop-out" of the target brand and inhibition of distractors.

Inspection of Table 1 shows that the FOA in the localization state is also guided by systematic search strategies, with the left–right and right–left strategies having large and similar posterior medians of .47 and .51. These results demonstrate guidance of attention by the scene's organization (Monk 1984; Ponsoda et al. 1995), and are evidence for the combined use of systematic and saliency-based search.

5.2 Switching Between Attention States

The estimates for refixations differ strongly and systematically between the two attention states. The posterior medians for refixations on the brand (.26) and text elements (−.25) in the localization state reveal that the FOA during localization is redirected somewhat to the brands, but away from the text. This is in strong contrast with the posterior median estimates for refixations on the brands (3.33) and text (2.55) in the identification state, both of which are much higher and positive, indicating that resampling package information, particularly its textual information, is important to verify the candidate's identity. The large estimates of the standard deviations of the heterogeneity distributions of these parameters suggests that whereas some individuals rely more on textual, others focus more on pictorial information such as the logo.

Table 3 presents the switching probabilities between the two unobserved attention states. It reveals the importance of both states in target search, corroborating the findings of Liechty et al. (2003) for scene perception tasks. It shows that participants are overall more likely to be in the localization state (median probability, .67) than in the identification state (.33). But still, one-third of the eye fixations are used to reduce uncertainty about brand identities. In line with this, there has been a growing recognition of the importance of object identification in computational approaches to visual search (Itti and Koch 2001). The probability of switching from localization to identification is much lower (median, .38) than the reverse (median, .78). To illustrate attention switching during target search over time, the estimated switching patterns of two participants are shown in the right part of Figure 3, along with their observed pattern of eye movements and estimated saliency maps in the left part. It illustrates the large individual differences in the saliency maps and in attention switching over time.

Table 4 shows that median fixation durations in the localization state (241 ms) are significantly longer than those in the identification state (213 ms). This difference is in accordance with research that fixation durations during scene perception (localization) are substantially longer than during reading (identification) (Rayner 1998), and is consistent with our findings about refixations in the two attention states.

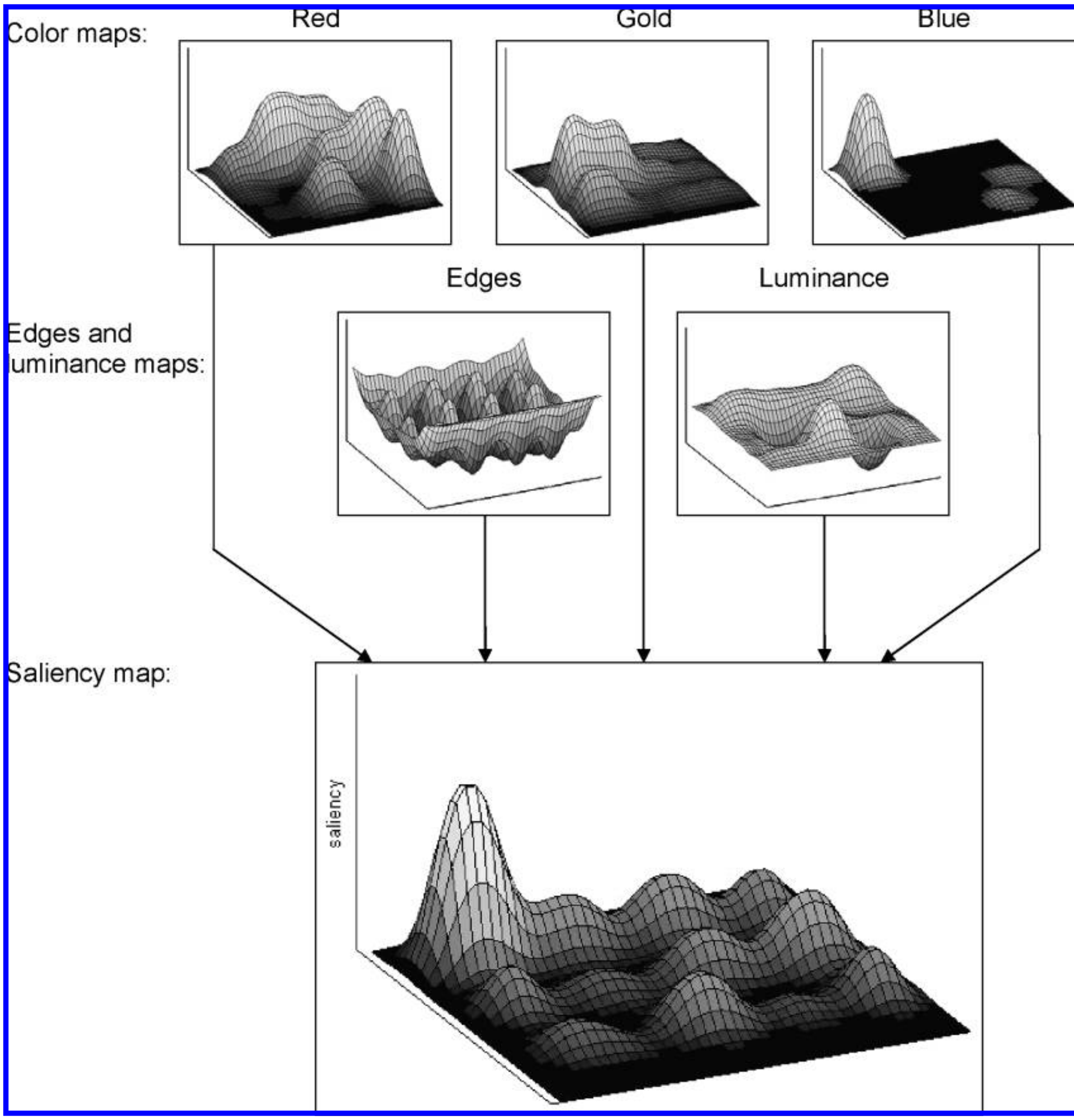


Figure 2. Estimated aggregate feature and saliency maps.

5.3 Exploring Effectiveness of Search Strategies

To gain insight into the model's external validity and evaluate the effectiveness of search strategies, we examine how individual-level parameter estimates (θ_{jhk}) relate to search accuracy and latency. We do this post hoc, with logistic regression analysis of the binary accuracy indicator and regression analysis of the log of search time (both variables were not used in model calibration). Eighty of the 106 participants found the target brand within the available 10 seconds; 14 participants found the wrong brand, and the remaining 12 ran out of time.

Table 3. Transition probabilities between localization and identification state of attention

From row to column	Localization state percentiles			Identification state percentiles		
Localization state	.57	.62	.67	.33	.38	.43
Identification state	.71	.78	.86	.14	.22	.29
Limiting probabilities	.63	.67	.71	.29	.33	.37

The individual-level parameter estimates predict search performance well (R^2 accuracy = .72; R^2 latency = .38) and remarkably consistently. That is, individuals placing a high weight on the positively diagnostic color blue found the target significantly faster ($\beta = -2.41$; standard error [SE] = .54) and more accurately ($\beta = 15.52$; SE = 6.37). Individuals directing their FOA toward the nondiagnostic color red, shared by many brands, were clearly less accurate ($\beta = -29.26$; SE = 7.52). Participants directing attention away from brighter regions were generally more accurate ($\beta = -22.93$; SE = 9.57). Directing attention to the negatively diagnostic color gold did not influence search performance, nor did emphasizing systematic search strategies. Time spent in the identification state negatively affected search accuracy ($\beta = -22.62$; SE = 8.52), but not latency of search. This is probably due to the fact that participants with only a weak template of the target brand need to spend more effort on identification. Individuals who refixate more on text information in the identification state (i.e., reading the labels on the package) tend to be more accurate ($\beta = 7.80$; SE = 4.71).

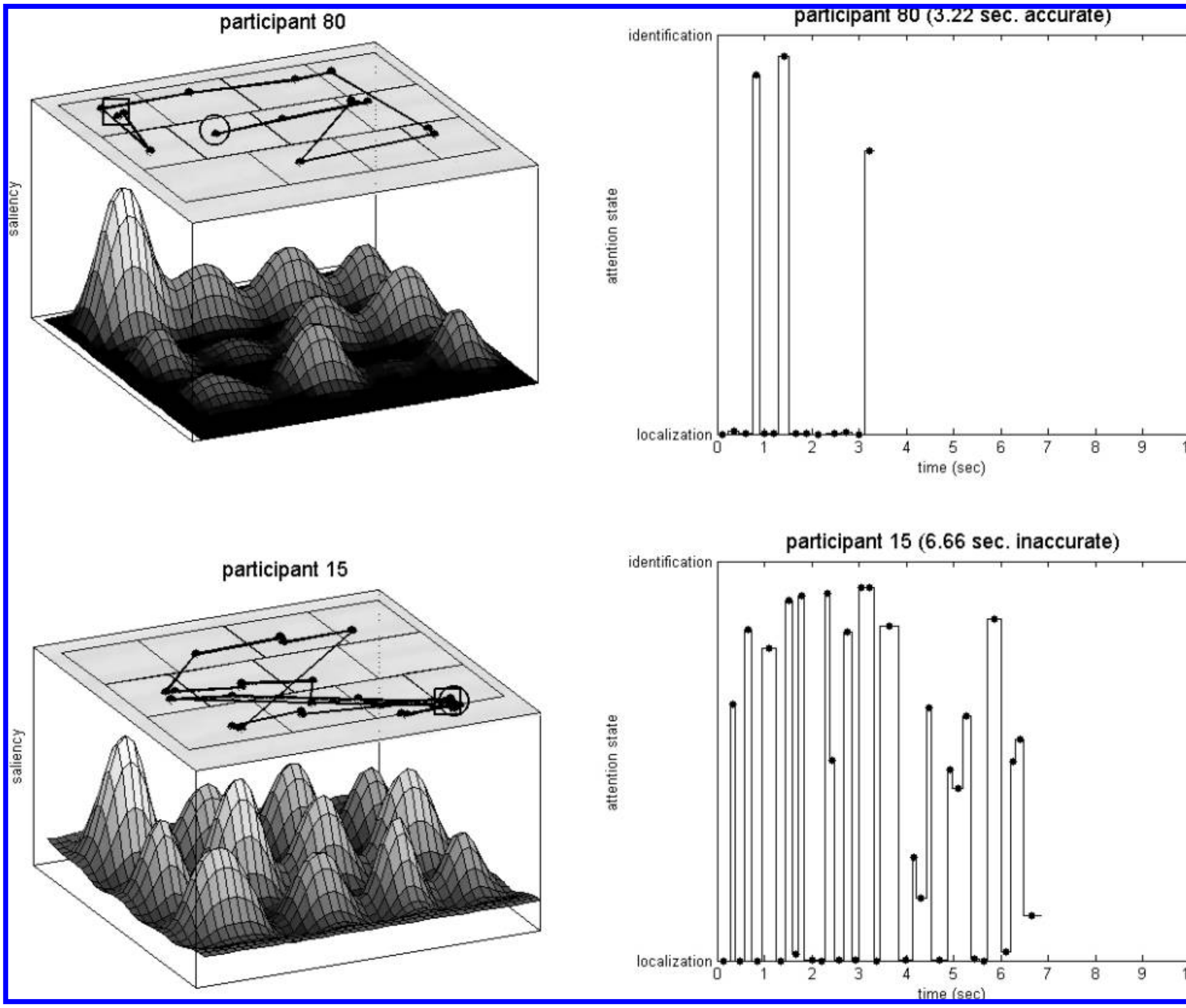


Figure 3. Switching between localization and identification and estimated saliency maps of two participants. Participant 80 (top panels) found the target in 3.22 seconds. Participant 15 (bottom panels) inaccurately indicated a distractor in 6.66 seconds. The estimated saliency maps (left), in which saliency is plotted vertically on the r–c coordinates of the display, demonstrate that the target stands out for participant 80 but not for participant 15. Compared with participant 15, participant 80's eye movements express a systematic pattern (a circle denotes the first, and a square the last eye fixation). The graphs with switching probabilities between the two attention states over time (right) show more intense switching for participant 15 compared with participant 80.

6. DISCUSSION

Eye tracking provides a powerful tool for better understanding visual search on complex displays, especially with modern technology that enables data to be collected at unprecedented scales, in academic and commercial settings. But identifying latent visual attention from eye-tracking data requires a statistical model such as the one presented here to allow inference on these key attention processes from eye movements. Our model (calibrated to eye movements) provides estimates of the feature maps and the resulting saliency map, thought to be maintained in specialized areas of the brain. It offers measures of the visual saliency of complex objects, describes the time course of the localization and identification states of attention, and determines the contribution of systematic search. Our results support the notion that people pre-consciously, but effectively, segment the scene and use the spatial arrangement of the segments to facilitate systematic search behavior; however, the use of such systematic search strategies was shown not to improve search performance. The results reflect the complexity of the task (Duncan and Humphreys 1989), requiring people to repeatedly switch to the identification state to determine the identity of a candidate and then switch back to continue localizing new candidates when the earlier one is not the target. Refixations on text are important for identification, and this reading improves the effectiveness of search.

Table 4. Mean fixation durations in localization and identification state of attention

	Mean fixation duration (ms) percentiles		
	.025	.500	.975
Localization state	238	241	245
Identification state	207	213	220

These findings reveal how individuals break up the complex task of target search into simpler subtasks between which they repeatedly switch to optimize search performance over time. The high spatiotemporal resolution of eye-tracking data allows our model to identify the activity of the attention states over time, presumably reflecting activity of the "where" and "what" pathways in the visual brain. An important area for future research is the further validation of this model component and its relation to these pathways. Other avenues for further research are to extend the model to include the duration of fixation, the role of peripheral vision, and indicators of search performance.

We believe that the application of the proposed model allows assessment of the effectiveness of search strategies and may aid in the optimization of search displays and the training of search agents. It may contribute to research in industrial engineering and human factors that seeks to uncover efficient search strategies to improve search performance. Applied to analyze the eye movements of experts, it may be used to develop guidelines for the training of novices (Wang et al. 1997) and to support the design of robotic vision systems, in which HMMs have already proven effective (Rimey and Brown 1991). Analyses such as the one provided in our empirical application can be used to optimize the design of search displays based on their estimated saliency. As a case in point, the model was used here to assess the saliency of brands, which influences in-store choices made by consumers (Drèze et al. 1994) and may be used for optimal package and shelf design.

Our results may also contribute to the further development of computational theories of visual search (Niebur and Koch 1998). In many cases, such computational models can incorporate more behavioral detail than a statistical model, not being hampered by the need to estimate the model parameters from behavioral data. But, similar to these models, our model is rooted in the biological architecture of the visual brain, whereas its estimation is facilitated by MCMC methodology. Whereas computational models reproduce experimental findings of visual search using simulation based on a priori defined parameter settings, the proposed model provides the crucial parameter estimates directly from eye movements observed during visual search. Thus our methodology may serve to obtain the parameter inputs of computational models and can complement these models by providing an empirically grounded understanding of visual search.

[Received August 2004. Revised May 2006.]

APPENDIX: PARAMETER ESTIMATES ON SYNTHETIC DATA

Parameter	True	Percentile .025	Percentile .500	Percentile .975	True	Percentile .025	Percentile .500	Percentile .975
	μ				σ			
Localization state								
Refixation brand	.20	−.19	.13	.40	.20	.13	.20	.31
Refixation text	−.20	−.35	−.17	−.02	.20	.13	.19	.29
Saliency								
Color								
Red	.20	.03	.14	.25	.20	.14	.20	.27
Gold	.00	−.11	−.00	.11	.20	.12	.17	.25
Blue	.30	.22	.28	.35	.20	.14	.18	.23
Luminence	−.10	−.21	−.13	−.05	.18	.12	.17	.24
Edges								
Brand group	−.20	−.26	−.18	−.10	.12	.11	.15	.20
Shelf	.50	.29	.43	.58	.15	.13	.19	.27
Systematic								
Horizontal zigzag								
Left–right	.50	.42	.50	.58	.20	.13	.18	.27
Right–left	.30	.14	.22	.30	.20	.14	.19	.29
Identification state								
Refixation brand	3.00	2.97	3.85	5.64	.30	.15	.26	.51
Refixation text	2.00	1.80	2.43	3.57	.30	.16	.28	.57
Attention switching								
	$\Pi_{.1}$				$\Pi_{.2}$			
$\Pi_{1.}$	.40	.36	.41	.47	.60	.52	.59	.64
$\Pi_{2.}$	.70	.60	.67	.78	.30	.22	.33	.40

REFERENCES

Anstis, S. M. (1974), "A Chart Demonstrating Variations in Acuity With Retinal Position," *Vision Research*, 14, 589–592.

Bullier, J., Schall, J. D., and Morel, A. (1996), "Functional Streams in Occipito-Frontal Connections in the Monkey," *Behavioural Brain Research*, 76, 89–97.

Bülthoff, H. H., and van Veen, H. A. H. C. (2001), "Vision and Action in Virtual Environments: Modern Psychophysics in Spatial Cognition Research," in *Vision and Attention*, eds. M. Jenkin and L. Harris, New York: Springer, pp. 233–252.

Chib, S. (1995), "Marginal Likelihood From the Gibbs Output," *Journal of the American Statistical Association*, 90, 1313–1321.

——— (1996), "Calculating Posterior Distributions and Modal Estimates in Markov Mixture Models," *Journal of Econometrics*, 75, 79–97.

Cressie, N. A. C. (1993), *Statistics for Spatial Data* (rev. ed.), New York: Wiley.

Damien, P., Wakefield, J., and Walker, S. (1999), "Gibbs Sampling for Bayesian Non-Conjugate and Hierarchical Models by Using Auxiliary Variables," *Journal of the Royal Statistical Society*, Ser. B, 61, 331–344.

Drèze, X., Hoch, S. J., and Purk, M. E. (1994), "Shelf Management and Space Allocation," *Journal of Retailing*, 70, 301–326.

Duchowski, A. T. (2003), *Eye Tracking Methodology: Theory and Practice*, London: Springer-Verlag.

Duncan, J., and Humphreys, G. W. (1989), "Visual Search and Stimulus Similarity," *Psychological Review*, 96, 433–458.

Findlay, J. M., and Gilchrist, I. D. (1998), "Eye Guidance and Visual Search," in *Eye Guidance in Reading and Scene Perception*, ed. G. Underwood, Amsterdam: Elsevier Science, pp. 295–312.

Friedman, A. (1979), "Framing Pictures: The Role of Knowledge in Automatized Encoding and Memory for Gist," *Journal of Experimental Psychology: General*, 108, 316–355.

Frühwirth-Schnatter, S. (2001), "Markov Chain Monte Carlo Estimation of Classical and Dynamic Switching and Mixture Models," *Journal of the American Statistical Association*, 96, 194–209.

Fuster, J. M. (2003), *Cortex and Mind: Unifying Cognition*, Oxford, U.K.: Oxford University Press.

Gevers, T. (2001), "Color-Based Retrieval," in *Principles of Visual Information Retrieval*, ed. M. S. Lew, London: Springer-Verlag, pp. 11–50.

Gramopadhye, A. K., Drury, C. G., and Sharit, J. (1997), "Feedback Strategies for Visual Search in Airframe Structural Inspection," *International Journal of Industrial Ergonomics*, 19, 333–344.

Henderson, J. M., Weeks, P. A. Jr., and Hollingworth, A. (1999), "The Effects of Semantic Consistency on Eye Movements During Complex Scene Viewing," *Journal of Experimental Psychology: Human Perception and Performance*, 25, 210–228.

Ho, G., Scialfa, C. T., Caird, J. K., and Graw, T. (2001), "Visual Search for Traffic Signs: The Effects of Clutter, Luminance, and Aging," *Human Factors*, 43, 194–207.

Horowitz, T. S., and Wolfe, J. M. (2003), "Memory for Rejected Distractors in Visual Search?" *Visual Cognition*, 10, 257–298.

Hoyer, W. D. (1984), "An Examination of Consumer Decision Making for a Common Repeat Purchase Product," *Journal of Consumer Research*, 11, 822–829.

Hubel, D. H., and Wiesel, T. N. (1962), "Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex," *Journal of Physiology*, 160, 106–154.

Inhoff, A. W., and Radach, R. (1998), "Definition and Computation of Oculomotor Measures in the Study of Cognitive Processes," in *Eye Guidance in Reading and Scene Perception*, ed. G. Underwood, Amsterdam: Elsevier Science, pp. 29–53.

Itti, L., and Koch, C. (2001), "Computational Modelling of Visual Attention," *Nature Reviews Neuroscience*, 2, 194–203.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., and Eastwood, J. D. (2003), "Attention Researchers! It Is Time to Look at the Real World," *Current Directions in Psychological Science*, 12, 176–180.

Liechty, J., Pieters, R., and Wedel, M. (2003), "Global and Local Covert Visual Attention: Evidence From a Bayesian Hidden Markov Model," *Psychometrika*, 68, 519–541.

Marr, D. (1982), *Vision: A Computation Investigation Into the Human Representation and Processing of Visual Information*, San Francisco: W.H. Freeman.

McCarley, J. S., Kramer, A. F., Wickens, C. D., Vidoni, E. D., and Boot, W. R. (2004), "Visual Skills in Airport-Security Screening," *Psychological Science*, 15, 302–306.

Monk, T. H. (1984), "Search," in *Sustained Attention in Human Performance*, ed. J. S. Warm, New York: Wiley, pp. 293–321.

Motter, B. C., and Holsapple, J. W. (2001), "Separating Attention From Chance in Active Visual Search," in *Visual Attention and Cortical Circuits*, eds. J. Braun, C. Koch, and J. L. Davis, Cambridge, MA: MIT Press, pp. 159–175.

Neal, R. M. (2003), "Slice Sampling," *The Annals of Statistics*, 31, 705–741.

Niebur, E., and Koch, C. (1998), "Computational Architectures for Attention," in *The Attentive Brain,* ed. R. Parasuraman, Cambridge, MA: MIT Press, pp. 163–186.

Pashler, H. E. (1998), *The Psychology of Attention*, Cambridge, MA: MIT Press.

Pashler, H. E., Johnston, J. C., and Ruthruff, E. (2001), "Attention and Performance," *Annual Review of Psychology*, 52, 629–651.

Pieters, R., and Wedel, M. (2004), "Attention Capture and Transfer in Advertising: Brand, Pictorial and Text Size Effects," *Journal of Marketing*, 68, 36–50.

Pomplun, M., Reingold, E. M., and Shen, J. (2003), "Area Activation: A Computational Model of Saccadic Selectivity in Visual Search," *Cognitive Science*, 27, 299–312.

Pomplun, M., Reingold, E. M., Shen, J., and Williams, D. E. (2000), "The Area Activation Model of Saccadic Selectivity in Visual Search," in *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society*, eds. L. R. Gleitman and A. K. Joshi, pp. 543–548.

Ponsoda, V., Scott, D., and Findlay, J. M. (1995), "A Probability Vector and Transition Matrix Analysis of Eye Movements During Visual Search," *Acta Psychologica*, 88, 167–185.

Rayner, K. (1998), "Eye Movements in Reading and Information Processing: 20 Years of Research," *Psychological Bulletin*, 124, 372–422.

Rimey, R. D., and Brown, C. M. (1991), "Controlling Eye Movements With Hidden Markov Models," *International Journal of Computer Vision*, 7, 47–65.

Robert, C. P. (1995), "Simulation of Truncated Normal Variables," *Statistics and Computing*, 5, 121–125.

Robert, C. P., Celeux, G., and Diebolt, J. (1993), "Bayesian Estimation of Hidden Markov Chains: A Stochastic Implementation," *Statistics and Probability Letters*, 16, 77–83.

Sanders, A. F., and Donk, M. (1996), "Visual Search," in *Handbook of Perception and Action: Attention 3*, eds. A. F. Sanders and O. Neumann, London: Academic Press, pp. 43–77.

Scott, S. L. (2002), "Bayesian Methods for Hidden Markov Models: Recursive Computing in the 21st Century," *Journal of the American Statistical Association*, 97, 337–351.

Thompson, K. G. (2005), "Dissociation of Selection From Saccade Programming," in *Neurobiology of Attention*, eds. L. Itti, G. Rees, and J. K. Tsotsos, San Diego: Elsevier, pp. 124–129.

Treisman, A., and Gelade, G. (1980), "A Feature-Integration Theory of Attention," *Cognitive Psychology*, 12, 97–136.

Ungerleider, L. G., and Mishkin, M. (1982), "Two Cortical Visual Systems," in *Analysis of Visual Behavior*, eds. D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield, Cambridge, MA: MIT Press, pp. 549–586.

Vora, J., Nair, S., Gramopadhye, A. K., Duchowski, A. T., Melloy, B. J., and Kanki, B. (2002), "Using Virtual Reality Technology for Aircraft Visual Inspection Training: Presence and Comparison Studies," *Applied Ergonomics*, 33, 559–570.

Wang, M.-J. J., Lin, S.-C., and Drury, C. G. (1997), "Training for Strategy in Visual Search," *International Journal of Industrial Ergonomics*, 20, 101–108.

Wedel, M., and Pieters, R. (2000), "Eye Fixations on Advertisements and Memory for Brands: A Model and Findings," *Marketing Science*, 19, 297–312.

Wertheimer, M. (2001), "Laws of Organization in Perceptual Forms," in *Visual Perception: Essential Readings*, ed. S. Yantis, Philadelphia: Psychology Press, pp. 216–224.

Wolfe, J. M. (1998), "Visual Search," in *Attention*, ed. H. Pashler, Hove, East Sussex, U.K.: University College London Press, pp. 13–73.

Yang, G.-Z., Dempere-Marco, L., Hu, X.-P., and Rowe, A. (2002), "Visual Search: Psychophysical Models and Practical Applications," *Image and Vision Computing*, 20, 291–305.

Zeki, S. (1993), *A Vision of the Brain*, Oxford, U.K.: Blackwell.

Zelinsky, G. J. (1996), "Using Eye Saccades to Assess the Selectivity of Search Movements," *Vision Research*, 36, 2177–2187.