Eye Fixations on Advertisements and Memory for Brands: A Model and Findings

Michel Wedel • Rik Pieters
Faculty of Economics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands, m.wedel@eco.rug.nl and University of Michigan School of Business, Ann Arbor, Michigan 48109-1234, wedel@umich.edu
Faculty of Economics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, F.G.M.Pieters@kub.nl

Abstract
The number of brands in the marketplace has vastly increased in the 1980s and 1990s, and the amount of money spent on advertising has run parallel. Print advertising is a major communication instrument for advertisers, but print media have become cluttered with advertisements for brands. Therefore, it has become difficult to attract and keep consumers' attention. Advertisements that fail to gain and retain consumers' attention cannot be effective, but attention is not sufficient: Advertising needs to leave durable traces of brands in memory. Eye movements are eminent indicators of visual attention. However, what is currently missing in eye movement research is a serious account of the processing that takes place to store information in long-term memory. We attempt to provide such an account through the development of a formal model. We model the process by which eye fixations on print advertisements lead to memory for the advertised brands, using a hierarchical Bayesian model, but, rather than postulating such a model as a mere data-analysis tool, we derive it from substantive theory on attention and memory. The model is calibrated to eye-movement data that are collected during exposure of subjects to ads in magazines, and subsequent recognition of the brand in a perceptual memory task. During exposure to the ads we record the frequencies of fixations on three ad elements; brand, pictorial and text and, during the memory task, the accuracy and latency of memory. Thus, the available data for each subject consist of the frequency of fixations on the ad elements and the accuracy and the latency of memory. The model that we develop is grounded in attention and memory theory and describes information extraction and accumulation during ad exposure and their effect on the accuracy and latency of brand memory. In formulating it, we assume that subjects have different eye-fixation rates for the different ad elements, because of which a negative binomial model of fixation frequency arises, and we specify the influence of the size of the ad elements. It is assumed that the number of fixations, not their duration, is related to the amount of information a consumer extracts from an ad. The information chunks extracted at each fixation are assumed to be random, varying across ads and consumers, and are estimated from the observed data. The accumulation of information across multiple fixations to the ad elements in long-term memory is assumed to be additive. The total amount of accumulated information that is not directly observed but estimated using our model influences both the accuracy and latency of subsequent brand memory. Accurate memory is assumed to occur when the accumulated information exceeds a threshold that varies randomly across ads and consumers in a binary probit-type of model component. The effect of two media-planning variables, the ad's serial position in a magazine and the ad's location on the double page, on the brand memory threshold are specified. We formulate hypotheses on the effects of ad element surface, serial position, and location.

The model is applied in a study involving a sample of 88 consumers who were exposed to 65 print ads appearing in their natural context in two magazines. The frequency of eye fixations was recorded for each consumer and advertisement with infrared eye-tracking methodology. In a subsequent indirect memory task, consumers identified the brands from pixelated images of the ads. Across the two magazines, fixations to the pictorial and the brand systematically promote accurate brand memory, but text fixations do not. Brand surface has a particularly prominent effect. The more information is extracted from an ad during fixations, the shorter the latency of brand memory is. We find a systematic recency effect: When subjects are exposed to an ad later, they tend to identify it better. In addition, there is a small primacy effect. The effect of the ad's location on the right or left of the page depends on the advertising context. We show how the model supports advertising planning and testing and offer recommendations for further research on the effectiveness of brand communication. In future research the model may be extended to accommodate the effects of repeated exposure to ads, to further detail the representation of strength and association of memory, and to include the effects of creative tactics and media planning variables beyond the ones we included in the present study.
(*Brand Advertising; Visual Attention; Brand Memory; Hierarchical Bayes*)

Marketing Science © 2000 INFORMS
Vol. 19, No. 4, Fall 2000, pp. 297–312
0732-2399/00/1904/0297/$05.00
1526-548X electronic ISSN

1. Introduction

> The first stage in the mental stream of thought leading to a purchase is attention. Thus, the first function of an advertisement is to attract attention. (C.H. Sandage 1945)

The number of brands in the marketplace has vastly increased in the 1980s and 1990s, and the amount of money spent on advertising has run parallel. Print advertising is a major communication instrument for advertisers, with estimated spending of over 30 billion dollars in 1995 in the United States alone. However, print media have become cluttered with advertisements for brands: many magazines have half of their pages or more carrying advertisements (Batra et al. 1996, Ha and Litman 1997). It has become difficult to attract consumers' attention in these cluttered media environments. Advertisements that fail to gain consumers' attention cannot be effective, but attention is not sufficient: Advertising needs to leave durable traces of brands in memory. It is a challenge for advertisers to rise high above the clutter of competing ads and to gain the attention of potential customers to build and sustain brand awareness. To achieve this goal, advertisers need to understand how consumers pay attention to print advertisements and how that contributes to memory for the advertised brands.

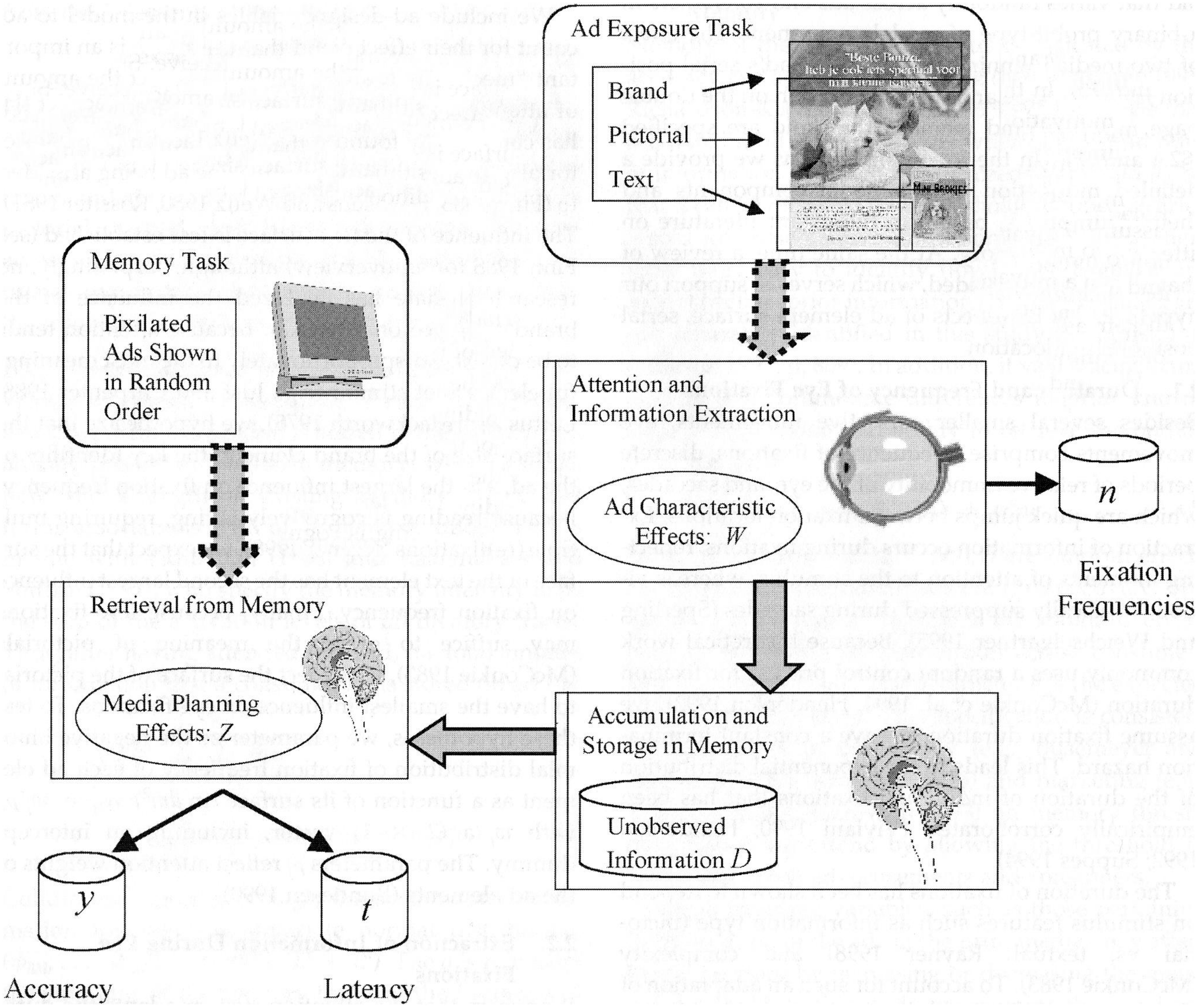
However, despite its importance to advertising effectiveness and "despite the tremendous amount of money spent on buying consumer attention, little to no research is done on attention" (Janiszewski and Bickart 1994, p. 329). Only recently a number of studies in marketing have aimed to fill the void in knowledge of attention processes (e.g., Janiszewski 1990a, 1990b, 1998; Olney et al. 1991; Rosbergen et al. 1997; Pieters et al. 1999). Our study contributes to this emerging stream of literature in several ways. First, following up on studies by Rayner (1998) and Suppes (1994), we propose a formal model describing how attention processes promote memory for the advertised brands. Specifically, we model the eye-fixations to advertisements and their effect on subsequent memory for the brand. The model is explicitly derived from recent developments in the theory of visual attention and memory. We accommodate the influence of ad-design variables (for example, the size of key ad elements, brand, pictorial, and text) on attention and of media planning factors, and of the ad's serial position and page location on memory performance. We test the model using disaggregate measures of attention and memory for a large sample of advertisements. Eye fixations on print ad elements and the subsequent accuracy and speed of memory for the advertised brands are assessed in a sample of 88 consumers and 65 print advertisements. Infrared eye-tracking methodology is used to record eye fixations and a perceptual identification task is used to assess brand memory. We investigate the influence for a large sample of advertisements, which is unique. Finally, we adopt a Bayesian approach to estimate the model and extensively investigate the model's assumptions.

2. A Model of Eye-Fixations Effects on Memory for Brands

According to Suppes (1994), what is missing in eye-movement research is a serious account of the processing that takes place to store information in long-term memory. In the remainder we attempt to provide such an account through the development of a formal model. The model is a hierarchical Bayesian model (see Gelman et al. 1995, Gilks et al. 1996), but, rather than postulating such a model as a mere data-analysis tool, we derive it from substantive theory on attention and memory, and we test its assumptions after it has been estimated. Although we are still somewhat remote from a true microlevel process model of attention and memory processes that is estimable from actual data, we believe we make a first step towards its development.

Before the model components are described in detail, we provide a brief overview that may serve to keep track of the sections that follow. Figure 1 provides a flow diagram of the process and collected data. The model is calibrated to eye-movement data that are collected during experimental exposure of 88 subjects to 65 ads, in two magazines, and subsequent recognition of the brand in a perceptual memory task (described in §3). During exposure to the ads, we record the frequencies of fixations on three ad elements: brand, pictorial, and text; and, during the memory task, the accuracy and latency of memory. Thus, the available data

Figure 1 Flow Diagram of Attention and Memory Processes in This Study.



for each subject consist of the numbers of fixations on the ad elements (denoted by n), and the accuracy (denoted by y), and the latency (denoted by t) of memory. In formulating the model, we assume that subjects have different eye-fixation rates for the different ad elements (because of which a negative binomial model of fixation frequency arises), and specify the influence of the size of the ad elements (§2.1). It is assumed (§2.2) that the number of fixations, not their duration, is related to the amount of information a consumer extracts from an ad. The information chunks extracted at each fixation are assumed to be random, varying across ads and consumers, and are estimated from the observed data. The accumulation of information across multiple fixations to the ad elements in long-term memory is assumed to be additive (§2.3). The total amount of accumulated information (denoted by D) that is not directly observed, but estimated using our model, influences both the accuracy and latency of subsequent brand memory. Accurate memory is assumed to occur when the accumulated information exceeds a threshold that varies randomly across ads and consumers in a binary probit-type of model component. The effect of two media-planning variables, the ad's serial position in a magazine and the ad's location on the double page, on the brand memory threshold are specified (§2.4 and 2.5). In the following sections we provide a detailed motivation of the model components and their assumptions, based on the extant literature on attention and memory. At the same time, a review of that literature is provided, which serves to support our hypotheses of the effects of ad element surface, serial position, and location.

2.1. Duration and Frequency of Eye Fixations

Besides several smaller corrective movements, eye movements comprise a sequence of fixations, discrete periods of relative immobility of the eye, and saccades, which are quick jumps between fixation locations. Extraction of information occurs during fixations, reflecting moments of attention to the stimulus, whereas vision is basically suppressed during saccades (Sperling and Weichselgartner 1995). Because theoretical work commonly uses a random control process for fixation duration (McConkie et al. 1994, Henderson 1992), we assume fixation duration to have a constant termination hazard. This leads to the exponential distribution of the duration of individual fixations that has been empirically corroborated (Viviani 1990; Henderson 1992; Suppes 1994).

The duration of fixations has been shown to depend on stimulus features such as information type (pictorial vs. textual, Rayner 1998) and complexity (McConkie 1983). To account for such an adaptation of the control process to differences between ad elements, we let the termination hazard of a fixation depend on ad element j. The total number of fixations, n_{ijl}, by consumer i at element j of ad l then follows a Poisson distribution. To account for the observed heterogeneity across consumers in fixation frequency to ad elements (Pieters et al. 1999), we let the parameter of the Poisson distribution follow a gamma distribution across consumers, which leads to the Negative Binomial Distribution for the fixation frequency.[1]

[1]The parameterization in the model is such that the expected number of fixations is $E(n_{ijl}) = \exp(\alpha_{1j})$.

$$n_{ijl} \mid \alpha_{1j}, \alpha_{2j} \sim NBD(\alpha_{1j}, \alpha_{2j}). \quad (1)$$

We include ad-designvariables in the model to account for their effect on attention. Surface is an important "mechanical" feature that may affect the amount of attention that ad elements receive. Starch research has consistently found that the surface size of the pictorial increases the likelihood of the ad being attended to (Finn 1988, Hanssens and Weitz 1980, Rossiter 1981). The influence of the text surface is less established (see Finn 1988 for an overview) although, surprisingly, no research to date has analyzed the influence of the brand's surface on attention. Because attention tends to be devoted disproportionately to the most meaningful elements of stimuli (e.g., Just and Carpenter 1988, Loftus and Mackworth 1978), we hypothesize that the surface size of the brand element, the key identifier of the ad, has the largest influence on fixation frequency. Because reading is cognitively taxing, requiring multiple (re)fixations (Rayner 1998), we expect that the surface of the text element has the second largest influence on fixation frequency. Finally, because few fixations may suffice to grasp the meaning of pictorials (McConkie 1983), we expect the surface of the pictorial to have the smallest influence on eye fixations. To test these hypotheses, we parameterize the negative binomial distribution of fixation frequency of each ad element as a function of its surface (in dm^2): $\alpha_{1j} = w_j'\rho_j$, with w_j a (2×1) vector, including an intercept dummy. The parameters ρ_j reflect attention weights of the ad elements (Bundesen 1990)

2.2. Extraction of Information During Eye Fixations

It appears that the attention system adapts the duration of fixations to local features of the stimulus to keep an approximately equal information rate per fixation (McConkie et al. 1994). For instance, longer fixations during reading mainly serve to improve the accuracy of the landing position of the next fixation (Coeffe and O'Regan 1987). According to Loftus (1972, p. 525–551) "... it might be assumed that information about a picture is transferred to LTS [the long-term memory store] in discrete chunks, with each chunk corresponding to a fixation." After such transfer, we assume that the information extracted per fixation is approximately constant. Let $k = 1, \ldots, n_{ijl}$ denote fixations. Then, the information extracted by subject i at the k-th fixation from the j elements of the l-th ad is:

$$\phi_{ijkl} = \mu_j + \sigma_j^a \zeta_l + \sigma_j^s \upsilon_i. \tag{2}$$

Here, μ_j is the expected quantity of information and $\zeta_l \sim N(0,1)$ and $\upsilon_i \sim N(0,1)$ are random effects for advertisement l and consumer i, respectively. The parameters σ_j^a and σ_j^s ($j = 1, \ldots, J$) represent the variability in pictorial, textual, and brand information extracted across ads and across consumers. The model thus postulates two independent sources of heterogeneity in the information extracted from the ad.[2]

2.3. Accumulation of Information in Memory

We assume that the information of each of the J types extracted during a fixation is added to the information already present in long-term memory, where it forms new and strengthens existing memory traces and forms associations with other memory traces. This is in line with Hintzman (1988) and Raaijmakers and Shiffrin (1992), who specify the memory intensity to be the sum of the activation levels of all memory traces. Thus, after having attended to an ad, the total amount of information that a consumer i has stored on ad l is:

$$\phi_{il} = \sum_{j=1}^{J} \sum_{k=1}^{n_{ijl}} \phi_{ijkl} = \sum_{j=1}^{J} (n_{ijl}\mu_j + n_{ijl}\sigma_j^a \zeta_l + n_{ijl}\sigma_j^s \upsilon_i). \tag{3}$$

Conditional upon $n_l = (n_{ijl})$, the accumulated information follows a multivariate normal distribution: $(\phi_{i1kl}, \ldots, \phi_{iJkl})' \sim MVN(\mu, \Sigma^s + \Sigma^a)$. The μ_j's represent the strengths of the memory traces of the J-types, and $\Sigma^s = diag(\sigma_j^{s^2})$ captures their variability across consumers. We assume the memory traces for the ad elements to be independent across consumers. The matrix Σ^a captures the variability across advertisements, where the off-diagonal elements relate to the associations between the memory traces of the J types. Note that our model captures the structure of long-term associative memory in a somewhat crude way, involving only the strengths and associations among J (brand, pictorial, text) types of memory nodes.

[2]The authors are grateful to Asim Ansari for suggesting this dual-heterogeneity representation.

2.4. Effect of Eye Fixations on Accuracy of Brand Memory

Memory of the advertised brands is indicated by the accuracy and speed of identifying them (LaBerge 1995, Richardson-Klavehn and Bjork 1988). We assume that accurate brand memory occurs when the total amount of information stored in LTM exceeds threshold θ_{il}, so that it can be retrieved. The threshold parameter varies across advertisements, because some brands are inherently more easy to identify (lower threshold) or are more familiar (prior information and the threshold are not separately identified in this study) (Hawley and Johnston 1991, p. 809). In addition, it varies across consumers to accommodate differences in prior knowledge or motivation and ability to retrieve the brands from memory:

$$\theta_{il} = \eta + \psi^s \vartheta_i + \psi^a \omega_l, \tag{4}$$

Here, $\vartheta_i \sim N(0,1)$ and $\omega_l \sim N(0,1)$ are random effects for consumer i and advertisement l, respectively. The parameters ψ^s and ψ^a represent the standard errors (s.e.) of the threshold distribution across consumers and ads, respectively. The parameter η is the expected value of the threshold. This specification is consistent with the work in psychology (e.g., Busemeyer and Townsend 1993, Kruschke 1996) and marketing (e.g., Krishnan and Chakravarti 1993) on memory thresholds, which we extend by allowing the threshold to vary across both advertisements and consumers.

Media planning variables, such as those pertaining to the location of the ads in the print media, may affect brand memory by increasing or decreasing the memory threshold. In this study we examine the serial position (front to back) and the page location (right or left) of advertisements, both of which are under marketing control and may influence brand memory (Rossiter and Percy 1997). Basic memory research consistently reports serial-position effects: primacy (first-to-last advantage) and recency (last-to-first advantage) (e.g., Greene 1986, Raaijmakers and Shiffrin 1992, Wyer and Srull 1986). However, serial-position findings from Starch scores are ambiguous (Finn 1988, Table 1). Starch research on the influence of page location also reports variously no, positive, and negative effects of the ad at the right (or left) page (Finn 1988, Table 1).

The present model allows tests of the influence of the ad's serial position and page location on the memory threshold. To account for primacy and recency effects simultaneously, we include both a linear and a quadratic term in serial position. A significant linear term identifies a primacy, if positive, or a recency effect, if negative. An additional quadratic term is expected to be positive when both primacy and recency effects are present. We model the threshold variation over ads ($\tilde{\theta}_l = \eta + \psi^a \omega_l$) as a function of the $q = 1, \ldots, Q$ media-planning variables, $\tilde{\theta}_l = z_l' \kappa + \nu \tau_l$, with $z_l = (z_{lq})$ and $\tau_l \sim N(0,1)$.

The memory indicator, $y_{il} = 1$ if brand l is accurately retrieved from memory, and $y_{il} = 0$ if not, is assumed to follow a Bernoulli distribution:

$$y_{il} \mid \pi_{il} \sim Binomial(1, \pi_{il}). \tag{5}$$

Accurate brand memory is assumed to occur with a probability:

$$\pi_{il} = \Pr(\phi_{il} + \epsilon_{il} > \theta_{il}) = \Pr(D_{il} > -\epsilon_{il}), \tag{6}$$

where D_{il} is implicitly defined and ϵ_{il} is the error occurring in memory retrieval (Hawley and Johnston 1991), assumed to follow a standard normal distribution. This leads to the probability of accurate memory retrieval of the brand to be $\pi_{il} = 1 - \Phi(D_{il})$, with $\Phi(\cdot)$ the cumulative normal distribution function. Note that because of the functional form of the normal CDF, the marginal effect of the additive accumulated information on accurate brand memory is decreasing, which is desirable (Kahneman 1973) and is in agreement with other models of memory retrieval (Bundesen 1990; Raaijmakers and Shiffrin 1992).

2.5. Effect of Eye Fixations on Latency of Brand Memory

Similar to memory accuracy, the speed of brand memory depends on the amount of information accumulated in memory (Sergent and Takane 1987, Richardson-Klavehn and Bjork 1988). We therefore also model latency of brand memory t_{il} as a function of the total amount of information accumulated D_{il}. Because the latency of brand memory is nonnegative, we model latency as:

$$t_{il} = \exp(\gamma_0 + \gamma_1 D_{il} + \delta_{il}), \tag{7}$$

with $\gamma = (\gamma_0, \gamma_1)'$ parameters and $\delta_{il} \mid \xi^2 \sim N(0, \xi^2)$ normally distributed error in retrieval, response, and recording, with ξ its s.e., so that the log latency is approximately constant, given the information retrieved.

2.6. Estimation and Checks of the Model

The model is estimated in a Bayesian framework, using an MCMC estimation algorithm (e.g., Gelman et al. 1995 and Gilks et al. 1996). In each iteration of the MCMC chain we draw from the full conditional distribution of parameter q, conditional on the values of the other parameters obtained from the previous draw: $\Xi_q \mid \Xi_{/q}, M$. Here $M = \{n, y, t\}$ denotes the data. The algorithm is described and tested in the Appendix.

Our model was built from a series of assumptions derived from attention and memory theory, summarized in Table 1. Because most of these assumptions have not yet been empirically tested, we subject our model to those tests. Rather than testing several nested and/or nonnested versions of the model to examine the adequacy of the model and test the assumptions, we choose the strategy to investigate whether the posterior distribution subsets of the parameters concentrate around zero, coupled with an extensive analysis of the model's residuals. An attractive feature of the MCMC methodology is that one can obtain accurate approximations of the small sample distributions of test statistics. These are computed from the Pearson residuals and fitted values from the parameter estimates $\Xi^{(r)}$ at each iterate r. We compute the residuals $e_n(\Xi^{(r)})$, $e_y(\Xi^{(r)})$, $e_t(\Xi^{(r)})$, and $e_\theta(\Xi^{(r)})$, as well as fitted values $\tilde{n}(\Xi^{(r)})$, $\tilde{y}(\Xi^{(r)})$, $\tilde{t}(\Xi^{(r)})$, and $\tilde{\theta}(\Xi^{(r)})$ for the eye fixations, memory accuracy, memory latency, and memory threshold. To test for misspecification of the model components, we compute the correlation between the residuals and fitted values: $\rho(e_n^{(r)}, \tilde{n}^{(r)})$, $\rho(e_y^{(r)}, \tilde{y}^{(r)})$, $\rho(e_t^{(r)}, \tilde{t}^{(r)})$, and $\rho(e_\theta^{(r)}, \theta^{(r)})$. In addition we compute the lag-one serial correlation (Judge et al. 1985, p. 399) of the memory residuals (pooled over subjects): $\rho(e_y^{(r)}, \Delta e_y^{(r)})$, $\rho(e_t^{(r)}, \Delta e_t^{(r)})$, and $\rho(e_\theta^{(r)}, \Delta e_\theta^{(r)})$ and of the total information $\rho(e_D^{(r)}, \Delta e_D^{(r)})$. Further normality of the information quantities is examined by computing the third, $m_3(\phi_j^{(r)})$, and fourth, $m_4(\phi_j^{(r)})$, moments of their distribution for each iteration. To investigate the independence of fixation frequency and information extracted[3] we compute the correlation of the information quantities with the fixation frequencies: $\rho(\phi_j^{(r)}, \tilde{n}_j^{(r)})$ for $j = 1, \ldots, J$ (Judge et al. 1985, p. 545). In Table 1 an overview of the test statistics is provided.

3. Study

3.1. Participants and Stimulus Material

A random sample of 88 female consumers between 19 and 52 years of age, were invited to the market research agency that collected the data. The stimulus material comprised all full-page advertisements inserted in two consumer magazines. The first magazine is an issue of a popular weekly magazine of a large retailer (*Allerhande*). It contains 108 pages, including 30 full-page advertisements. The ads promote food products (wine, tea, ice cream, eggs, coffee, meat, soup, dairy products, candy), personal care products (shampoo, bath foam, shaving foam, deodorant), cigarettes, and magazines. The second magazine is an issue of a glossy monthly magazine (*Cosmopolitan*) containing 128 pages and 35 full-page advertisements. The ads promote different personal care products (after-shave, nail polish, eau de toilette, body lotion), lingerie, cars, cigarettes, magazines, and food products (fruit juice, sauces, beer). Advertisements are located throughout the target magazines with different facing editorial material. All sixty-five ads contain the three key ad elements: brand, pictorial, and text.

[3]Note that the assumption of independence of the random coefficients and regressors is made in all hierarchical regression models but is rarely investigated.

3.2. Eye-Movement Recording

Participants were seated at a table on which the target magazine was fixed in such a way that it could be paged through freely. Participants engaged in a typical visual exploration task, as they would at home or in a waiting room (Janiszewski 1998). They received the magazines in randomized order. While paging at their own pace, participants' eye movements were recorded with infrared corneal reflection eye-tracking methodology (e.g., Ober 1994). It measures the position of the fovea at 50 Hz. The specific eye-tracking equipment we used has been developed by Verify International in Rotterdam (Netherlands). It allows participants to freely move their heads within a virtual box of about

Table 1 **Summary of the Model, Symbols, Assumptions, and Diagnostic Tests**

Section	Model Component	Symbol	Distribution	Assumptions	Diagnostic Statistics
2.1	Fixation frequency	n_{ijl}	$NBD(w'\rho_j, \alpha_{2j})$	Constant termination rate, consumer heterogeneity, linear surface effect.	$\rho(e_n^{(r)}, \tilde{n}^{(r)})$
2.2	Information extracted	ϕ_{ijl}	$N(\mu_j, \sigma_j^{a2} + \sigma_j^{s2})$	Information varies across ads and consumers, independent of fixations.	$\rho(\phi_j^{(r)}, \tilde{n}_j^{(r)})$
2.3	Information accumulated	ϕ_i	$MVN(n_i'\mu, n_i\Sigma n_i')$	Additive accumulation of normal distributed information.	$m_3(\phi_j^{(r)}), m_4(\phi_j^{(r)})$
2.4	Memory threshold	θ_{il}	$N(z_l'\kappa, \nu^2 + \psi^{s2})$	Ad and consumer heterogeneity, linear effects of media planning variables.	$\rho(e_\theta^{(r)}, \theta^{(r)}), \rho(e_\theta^{(r)}, \Delta e_\theta^{(r)})$
2.4	Total information	D_i	$MVN(\phi_i - \theta_i, I_L)$	Prior information combined with threshold, independent across ads.	$\rho(e_D^{(r)}, \Delta e_D^{(r)})$
2.4	Memory accuracy	y_i	$B(1 - \Phi(D_i))$	Diminishing information returns, independent across ads.	$\rho(e_y^{(r)}, \tilde{y}^{(r)}), \rho(e_y^{(r)}, \Delta e_y^{(r)})$
2.5	Memory latency	t_{il}	$LN(\gamma_0 + \gamma_1 D_{il}, \xi^2)$	Information influences accuracy and latency, independence across ads	$\rho(e_t^{(r)}, \tilde{t}^{(r)}), \rho(e_t^{(r)}, \Delta e_t^{(r)})$

Index: i = 1, . . . I, consumers; j = 1, . . . , J, ad elements; k = 1, . . . , K, eye fixations; l = 1, . . . , L, advertisements.

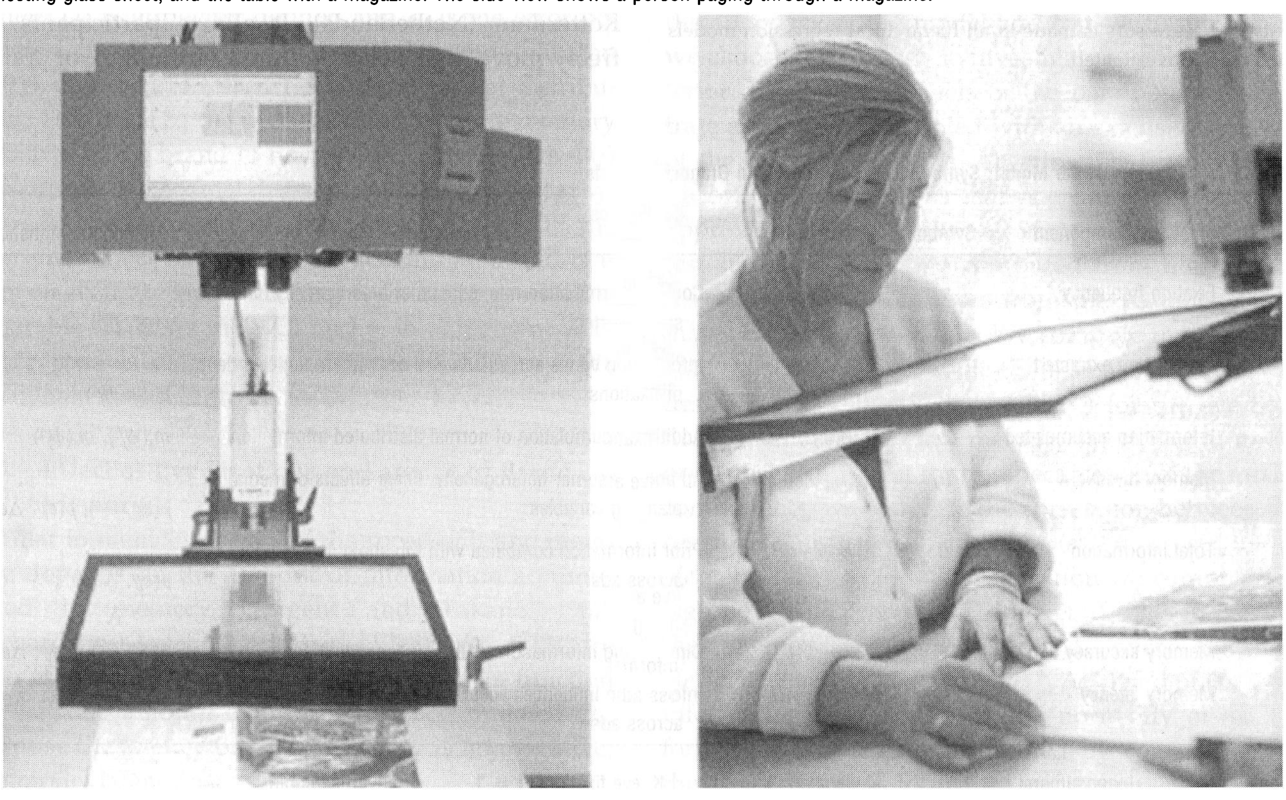
30 centimeters while cameras track the position of both the head and the eyes, and a computer continuously matches the information. Another camera records the pages to which the magazine is opened. The computer also matches the information of eye movements and magazine pages such that the exact locations of individual fixations on the magazine become available. For all sixty-five ads, fixation frequencies on the three ad elements—brand, pictorial and text—were recorded. We define the brand element of print advertisements as all pictorial and textual references to the brand, including the name, logo, symbols, and packshot. The pictorial element includes all pictorial information in the advertisement, such as illustrations, but excluding the brand symbols. The textual element includes all text in the advertisement, such as the headline, sublines, payoff, and body text but excludes references to the brand name. Figure 2 offers a front and a side view of the experimental setup.

The front view shows the infrared eye-tracking equipment, with the infrared and head location cameras and the control panel at the top, the page recognition camera and an adjustable glass sheet in the middle, and the table with a magazine at the bottom. The side view shows a person paging freely through a magazine. The glass sheet between the magazine and the person reflects the infrared light beam coming from the top onto the cornea, and from the cornea back to the top, but all other light passes. After eye-movement recording, participants engaged in another unrelated study, which took about 20 minutes and was followed by the surprise memory task.

3.3. Perceptual Memory Task

An indirect (perceptual) memory task was used to assess attention effects on brand memory. In an indirect memory task, participants engage in a cognitive or motor activity, with instructions referring only to the task

Figure 2 The Verity Eye-Movement Recording Equipment.
Note: The front view shows the infrared and head location cameras and the control panel, the page recognition camera and an infrared reflecting glass sheet, and the table with a magazine. The side view shows a person paging through a magazine.



without time- and location-retrieval cues, and without references to prior exposure to the stimuli (Richardson-Klavehn and Bjork 1988). To probe perceptual memory, a specific form of indirect memory, participants identify stimuli that are incomplete, presented briefly, or perceptually degraded (Hawley and Johnston 1991). The strength of perceptual memory then is assessed as the increased accuracy and/or speed with which participants identify the target stimuli after exposure.

There are several considerations that favored the use of an indirect memory task, specifically, a perceptual memory task, in this study. First, the goal of advertising is to promote memory for the advertised brands—indirect memory—instead of memory for the advertisements—direct memory—(Krishnan and Chakravarti 1993, 1999). Furthermore, indirect memory measures tend to have a lower threshold (Cowan 1995) and thus are more sensitive to the small effects of advertisements in highly competitive media (Jacoby and Dallas 1981). Also, in making daily decisions for low-involvement products, consumers tend to invest little effort in retrieving explicit information from advertisements and instead may rely more heavily on their indirect perceptual memory (cf. Krishnan and Shapiro 1996). The perceptual memory task was administered as follows. Participants were seated individually in front of an NEC 21-inch, touch-sensitive monitor to assess perceptual memory for brands. They received detailed instructions and examples of the task. On the monitor, the 65 advertisements from the two magazines were shown, as well as 10 other ads (mirroring the target ads in products and design), in random order. Pixelated images of the advertisements were shown (e.g., Hawley and Johnston 1991), in which colors and forms of pictorials and logos remain largely intact, and big type from brand name, headline, and subheadings can be identified. Each ad was accompanied by four brand names. Participants were asked to identify the target brand in the ad as soon as possible by touching the correct brand name on the screen. The distracting brand names were chosen to maximize the likelihood of incorrect choices in case of guessing. Accuracy of brand memory and response latency were recorded for each advertisement and consumer.

4. Findings

Table 2 presents a number of descriptive statistics for the data. On average, the pictorial has the largest surface, followed by text and then brand, for both magazines. *Allerhande* has half of its ads on the right page, which is somewhat more than that for *Cosmopolitan*. For both magazines, consumers fixate most frequently on the pictorial, followed by the text and then the brand. Accurate identification of the brands occurs in around half of the cases, which takes consumers 2.5 seconds on average. The model was calibrated to the data on the two magazines separately, which allows us to compare results across advertising vehicles. We report the findings by model component below.

4.1. Fixation Frequency on Ad Elements

Table 3 presents the median and 95% credible intervals of the posterior distribution of the fixation frequency parameters. As predicted, the largest effect of the surface of the ad elements is observed for the brand. The text receives less fixations per surface area, the pictorial the least. The disproportionately high unit-fixation frequencies on the brand relative to the other two ad elements is striking. This supports our hypothesis that the brand element is the most meaningful ad element. The posterior distributions of the fixation frequency parameters for the two magazines do not overlap. The brand surface effect is higher for *Allerhande* than for *Cosmopolitan*, while the reverse holds for the pictorial

Table 2 **Descriptive Statistics**

	Magazines			
	Allerhande		Cosmopolitan	
Variable	Mean	Variance	Mean	Variance
Surface brand (dm^2)	0.35	0.07	0.44	0.14
Pictorial (dm^2)	3.84	0.94	4.36	0.77
Text (dm^2)	1.94	0.85	1.20	1.23
Right-page advertisements	0.50	0.26	0.57	0.25
Fixation frequency brand (n_1)	1.78	7.25	1.71	7.43
Pictorial (n_2)	5.37	27.48	4.91	24.64
Text (n_3)	4.40	36.53	3.02	19.19
Memory accuracy (y)	0.44	0.25	0.51	0.25
Memory latency (t, sec)	2.44	1.08	2.42	1.02

Table 3 Parameter Estimates for Fixation Frequency: Median and 95% Credible Intervals

	Magazines					
	Allerhande			Cosmopolitan		
Parameter	0.025	0.500	0.975	0.025	0.500	0.975
Brand						
ρ_{01} (intercept)	−0.500	−0.401	−0.298	−0.269	−0.182	−0.097
ρ_{11} (surface)	2.045	2.249	2.475	1.184	1.334	1.472
α_{21}	−0.389	−0.290	−0.197	−0.506	−0.418	−0.330
Pictorial						
ρ_{02} (intercept)	0.825	0.978	1.132	−0.014	0.213	0.425
ρ_{12} (surface)	0.138	0.178	0.271	0.264	0.310	0.353
α_{22}	0.128	0.194	0.259	0.258	0.320	0.391
Text						
ρ_{03} (intercept)	0.741	0.876	1.023	0.288	0.387	0.483
ρ_{13} (surface)	0.229	0.293	0.358	0.448	0.509	0.573
α_{23}	−0.614	−0.544	−0.476	−0.661	−0.586	−0.513

and text. This may be explained by the editorial content of the magazines, which is much more focused on shopping and brand information for *Allerhande* (e.g., Pieters and Warlop 1999).

4.2. Accuracy of Brand Memory

Table 4 offers the median and 95% credible intervals of the posterior distributions of the information parameters. Brand memory for ads in *Allerhande* was less accurate than for ads in *Cosmopolitan* (as reflected in a higher value of the memory threshold parameter). The heterogeneity in brand memory is larger for *Allerhande* than for *Cosmopolitan,* and it is significantly larger across advertisements than across consumers. This indicates that differences in brand memory between advertisements are more prominent than differences between consumers.

Table 5 shows the parameter estimates of the effects of the media planning variables on memory threshold. There is strong evidence of a recency effect in serial position: The threshold decreases with an increasing page number, leading to better memory. The quadratic effect is positive for both magazines, but the 95% credible interval covers the zero value for *Allerhande*. This quadratic effect is indicative of a primacy effect as well

Table 4 Parameter Estimates for Accuracy of Brand Memory: Median and 95% Credible Intervals

	⟨Magazines⟩					
	Allerhande			Cosmopolitan		
Information Parameters	0.025	0.500	0.975	0.025	0.500	0.975
Thresholds						
η (mean)	0.856	0.467	0.080	0.474	0.191	−0.080
ψ^a (ad s.e.)	0.721	0.957	1.321	0.576	0.756	1.045
ψ^s (consumer s.e.)	0.308	0.431	0.573	0.285	0.386	0.497
Fixation frequency						
Brand						
μ_1 (mean)	0.013	0.057	0.107	0.021	0.062	0.104
σ_1^a (ad s.e.)	0.041	0.069	0.114	0.041	0.066	0.109
σ_1^s (consumer s.e.)	0.035	0.064	0.107	0.029	0.046	0.080
Pictorial						
μ_2 (mean)	−0.004	0.020	0.047	0.009	0.033	0.054
σ_2^a (ad s.e.)	0.031	0.045	0.068	0.028	0.039	0.058
σ_2^s (consumer s.e.)	0.028	0.041	0.057	0.026	0.037	0.052
Text						
μ_3 (mean)	−0.016	0.010	0.032	−0.029	0.003	0.032
σ_3^a (ad s.e.)	0.031	0.045	0.067	0.035	0.053	0.080
σ_3^s (consumer s.e.)	0.025	0.035	0.048	0.029	0.043	0.062

Table 5 Parameter Estimates for Memory Threshold: Median and 95% Credible Intervals

	⟨Magazines⟩					
	Allerhande			Cosmopolitan		
Threshold Parameter	0.025	0.500	0.975	0.025	0.500	0.975
κ_0 (mean)	1.441	1.134	0.827	1.030	0.800	0.592
κ_1 (ad page number × 100)	−1.008	−2.187	−3.334	−0.612	−1.411	−2.181
κ_2 (ad page number: squared)	2.051	0.973	−0.088	1.255	0.674	0.071
κ_3 (ad page location: right)	0.447	0.269	0.088	−0.091	−0.239	−0.396
ν (residual s.e.)	0.073	0.084	0.096	0.059	0.070	0.083

as a recency effect, although the effect of the latter is stronger and more consistent across the two magazines. The effect of page location reverses signs between the magazines: For *Allerhande* the positive parameter value indicates a better memory of left-page ads, the negative sign for *Cosmopolitan* indicates better memory of right-page ads.

For both magazines, the amount of information extracted on a single fixation decreases from the brand to pictorial to text (reflected in the posterior medians of μ_1, μ_2, and μ_3). Thus, although the brand has the smallest surface of the ad elements (Table 2), and although consumers fixate less frequently on the brand element (Table 2), it receives the most fixations per surface unit (Table 2) and, moreover, delivers most information for subsequent accurate brand memory. This underlines the importance of eye fixations on the brand element of advertisements in building brand memory. The posterior distribution of brand information is very similar across the two magazines and there is substantial heterogeneity across ads and consumers.

Although the pictorial elements do not contain the brand name, logo, or pack shot, eye fixations on the pictorial do deliver information relevant for later brand memory (μ_2). The amount of information extracted is half or less than that extracted by fixations on the brand but is still substantial (for *Allerhande* ads the pictorial contributes less information, and the 95% credible interval just covers the zero value). The amounts of pictorial information extracted vary considerably across ads and consumers. These findings underscore the signal value that pictorials have for brand memory, the differences in sensitivity of consumers for pictorial information, and the differences in quality of pictorials to contribute to brand memory.

Note that although the fine print in the body text tends to become illegible because of the ad pixilation in our memory task, the headline, subheadline and payoff remain legible in most cases. In addition, the block of body text is clearly recognizable, and its shape, color, and location provide contextual cues for brand memory (Bundesen 1990). Still, for both magazines, the 95% interval of the information extracted from the text element of the advertisements (μ_3) covers the zero value. Hence, eye fixations on the text elements of advertisements in this study, on average, do not yield information that promotes brand memory. The amount of variation of textual information across ads and consumers is fairly large and is comparable to that of the pictorial information, so that for some ads there may be a significant brand memory trace associated with textual information. The posterior distributions of the correlations among the threshold, brand, pictorial, and textual information, indicative of associations among those memory traces, all concentrate around the zero value (they are, therefore, not shown in the table).

4.3. Latency of Brand Memory

Table 6 presents the results for the latency of brand memory. The mean response time, adjusted for the information extracted during exposure to the ads, is quite similar to that for ads in *Allerhande* and in *Cosmopolitan*. As expected, the information effect of eye fixations on identification latency is negative in both studies. This shows that, as predicted, information accumulated in memory through eye fixations promotes *faster* identification of brands. The effect is significantly more prominent in *Cosmopolitan* than in *Allerhande*. Note that the effects of brand, pictorial, and text fixations on latency are proportional to those reported in the previous section.[4]

4.4. Diagnostic Checks

Because our model assumptions, derived from attention and memory theory, have mostly not been subjected to stringent empirical testing in previous research, we perform an extensive series of model checks, based on the posterior distribution of the test statistics that we described earlier. Table 1 provides an overview of the test statistics computed. The majority of the assumptions cannot be rejected based on the data: From the 38 residual tests performed, only five appeared "significant." For all ad elements the information extracted follows an approximate normal distribution and is not correlated with the total number of fixations, the total amount of information does not display serial correlation, and there is no evidence for

[4]The indirect effects of ad element information on latency can be obtained as $\gamma^1\mu^1$, $\gamma^1\mu^2$, and $\gamma^1\mu^3$.

Table 6 Parameter Estimates for Latency of Brand Memory: Median and 95% Credible Intervals

	Magazines					
	Allerhande			Cosmopolitan		
Information Parameter	0.025	0.500	0.975	0.025	0.500	0.975
γ_0 (mean)	0.724	0.746	0.769	0.755	0.776	0.796
γ_1 (information)	−0.076	−0.056	−0.035	−0.111	−0.085	−0.063
ξ (residual s.e.)	0.573	0.588	0.604	0.521	0.535	0.549

misspecification of the threshold and accuracy of memory, whereas the latter do not show serial correlation. However, memory latency, although not misspecified, does exhibit minor residual serial correlation, evidenced by $\rho(e_t^{(r)}, \Delta e_t^{(r)})$ (95% credible interval in parentheses), in both *Allerhande* (0.05, 0.06) and *Cosmopolitan* (0.12, 0.13). The pictorial fixation frequency residuals are positively correlated with the predicted values, $\rho(e_n^{(r)}, \tilde{n}^{(r)})$ in *Allerhande* (0.01, 0.06) and *Cosmopolitan* (0.02, 0.08). This indicates a slight nonlinear effect of the pictorial surface. Text fixations in *Cosmopolitan* show a similar effect (0.03, 0.07). Because these correlations are very low and the large number of tests may give rise to some chance capitalization, we have not modified the model and conclude that it is quite a reasonable description of the data.

5. Conclusion

Ever since Karslake (1940) used the Purdue Eye Camera to collect eye-movement data on advertisements appearing in the *Saturday Evening Post*, research in this area has been hampered by the lack of adequate models for the analysis of eye-tracking data. This has prevented the accumulation of generalizable findings on attention for advertising and its effects on brand memory. The model outlined here provides the first step towards a formal conceptualization and representation of attention processes that build memory for brands. Rather than developing a mere data-analysis tool, we have gone to great lengths to provide a representation of the processes of fixation and memory based on current theory on those processes, supporting each of the assumptions made from previous literature. The proposed model of attention and memory processes can facilitate future research in marketing, as each of its parameters has substantive meaning and can be conditioned on relevant ad, media, and consumer characteristics. Of course, the model is only a first step, and many of its aspects need further detail and future testing. It may be extended, for example, to accommodate the effects of repeated exposure to ads, to further detail the representation of strength and association of memory, and to include the effects of ad layout characteristics and media planning variables beyond the ones we included in the present study.

Starch scores have been fruitfully used for almost a century to support advertising and media planning decisions. Most academic research based on Starch scores has focused on the role of pictorial and textual, mechanical, and psycholinguistic features of advertisements. Unfortunately, Starch results are difficult to interpret because they are retrospective reports of attention, which confound attention and memory processes by asking consumers to remember how they previously paid attention to ads. Eye-tracking data allow proper tests of the attention effects of the surface size of ad elements and other ad design characteristics. Our findings based on such eye-tracking data reveal a surprising amount of information in the brand element of ads. On average, the surface of the brand element was much smaller than the surface of the pictorial (about 10 times) and the text (about 3–5 times) elements of the print advertisements. In addition, the fixation frequency on the brand element is much lower than the fixation frequencies on the pictorial and text elements. However, the brand element receives by far the most eye fixations *per unit of its surface*, followed by the text element and the pictorial. Although we predicted that the brand element would receive the most eye fixations per surface unit, the large differences with the text and pictorial element are striking. They show that even in an exploratory search context in which consumers freely page through magazines, and in which they dwell only for a short moment on each ad (less than 3 seconds), the brand element draws a disproportionately large amount of attention. These findings support marketing's longtime emphasis on the role of the brand in advertising. At the same time, they underline the incomplete and potentially erroneous results of previous advertising research that has ignored brand effects on attention and memory. Ours is the first study to identify the magnitude of the brand effect using direct measures of attention across a large sample of ads. The significant effect of fixations to the pictorial attests to the importance of visual cues in building brand memory. The nonsignificance of text effects on brand memory in this study may be partly intrinsic to the specific task used, but it also supports previous findings that memory for textual information is less than that for pictorial information. The use of explicit verbal cues in the memory task, such as words from the headline or the payoff, may result in larger effects of text fixations on brand memory, and future research may pursue this.

Combined, our findings on the importance of the brand element might stimulate marketers to greatly enlarge the brand surface of ads in order to maximize ad effectiveness. Several considerations warn against injudicious application of such a strategy. First, in our study most attention in an absolute sense went to the pictorial of the ads, and the textual element also received more attention than the brand element did. Second, the pictorial element had an effect not to be ignored, on brand memory. Also, the heterogeneity in the effect of the textual information indicates that for some ads, the text did contribute to brand memory. Third, ads need to gain attention and build brand memory across repeated exposures, and the ad's sensitivity to attention and memory wearout may well be higher when the brand surface is disproportionately large. In our sample of ads, the variation in brand element surface was limited. Whereas the ad sample was large and comprised ads across product categories from two different magazines, only with great caution should the effects of ad element surface be extrapolated beyond the range of values in this study.

The identified strong recency effect is consistent with the model and findings of Wyer and Srull (1986). They show recency effects under conditions of high information load, which we believe to be similar to what consumers were facing in our study. Under high information load, earlier stimuli tend to be displaced from short-term memory by later stimuli in the sequence, which lowers the likelihood of the earlier stimuli being stored and subsequently retrieved. Retroactive inhibition in long-term memory by later stimuli in the sequence may further contribute to retrieval difficulties of the earlier stimuli, which adds to the recency effect (e.g., Burke and Srull 1988). This suggests that advertisers who seek to maximize brand memory may want to place their ads towards the end of magazines. However, albeit smaller in magnitude and only significant in one of the two magazines, we also recovered a primacy effect. This primacy effect may be dependent upon the advertising context: the fact that it was not significant for *Allerhande* may have been caused by it being cluttered with (large and small) advertisements. Longer time intervals between exposure and memory retrieval, repeated exposures, and advertising contexts may modify the recency and primacy effects found in this study. The effect of page position on the memory threshold differed between magazines. A tentative explanation might be that a majority of full-page ads in *Allerhande* (80%) and only a minority of full-page ads in *Cosmopolitan* (25%) was faced by a page with multiple small ads. Most ads in *Cosmopolitan* were faced by editorial material. Perhaps location effects are positive when the left page contains editorial material but negative when the left page contains ads for competing brands. The result calls for further research into the effects of media context on advertising effectiveness. In the meantime, the recommendation of Rossiter and Percy (1997, Chapter 10) not to adjust systematically for page location in media planning seems warranted.

We are somewhat reluctant to formulate "universal" recommendations for print-ad execution, as our study has assessed attention and memory in highly controlled conditions. Rather, we think that ad strategy implications should be derived from empirical generalizations drawn from a larger number of studies involving more widely varying measures and conditions. Keeping these reservations in mind, we believe that our model and findings are pertinent to at least two crucial components of ad execution: copy and media strategy. First, the estimates of the effects of ad characteristics on attention and of the information quantities extracted are instrumental in the development of copy strategy. Our study suggests that the effects of ad element surfaces on attention are relatively

Table A1 Results of Synthetic Data Analyses

	Data Set 1				Data Set 2			
Coeff.	True	0.025	0.500	0.975	True	0.025	0.500	0.975
η	−1.000	−1.393	−1.186	−0.973	−1.000	−1.088	−0.972	−0.865
μ_1	0.100	0.050	0.091	0.131	0.100	0.081	0.128	0.182
μ_2	0.100	0.082	0.126	0.167	0.100	0.025	0.085	0.138
μ_3	0.100	0.068	0.108	0.143	0.100	0.026	0.079	0.125
ψ^a	0.100	0.077	0.098	0.127	0.100	0.039	0.075	0.167
σ_1^a	0.100	0.079	0.099	0.127	0.100	0.056	0.083	0.122
σ_2^a	0.100	0.081	0.100	0.129	0.100	0.071	0.095	0.132
σ_3^a	0.100	0.078	0.098	0.128	0.100	0.062	0.087	0.123
ψ^s	0.500	0.428	0.515	0.625	0.200	0.059	0.116	0.203
σ_1^s	0.000	0.000	0.000	0.000	0.200	0.126	0.160	0.204
σ_2^s	0.000	0.000	0.000	0.000	0.200	0.190	0.226	0.272
σ_3^s	0.000	0.000	0.000	0.000	0.200	0.134	0.168	0.208
γ_0	1.000	0.978	0.996	1.015	1.000	0.974	0.993	1.011
γ_1	−0.100	−0.123	−0.105	−0.088	−0.100	−0.133	−0.118	−0.105
ξ	0.500	0.486	0.498	0.511	0.500	0.477	0.488	0.503
$\exp(\alpha_{11})$	2.000	1.922	2.018	2.120	4.000	3.982	4.134	4.313
$\exp(\alpha_{21})$	0.500	0.451	0.491	0.533	1.000	0.915	0.970	1.044
$\exp(\alpha_{12})$	5.000	4.738	4.901	5.071	1.000	0.908	0.967	1.025
$\exp(\alpha_{22})$	2.000	1.795	1.936	2.040	0.500	0.456	0.500	0.553
$\exp(\alpha_{13})$	3.000	2.773	2.924	3.081	5.000	4.850	5.104	5.385
$\exp(\alpha_{23})$	0.500	0.477	0.511	0.540	0.500	0.467	0.496	0.530

invariable across advertising contexts, which may be an important asset in the development of advertising copy. In addition, the information extracted from the brand element was very similar across advertising media. It is remarkable that previous attention research, in particular the many studies based on Starch data, have taught us surprisingly little about the role of the brand, the element that is most informative to consumers and most important to advertisers and ad agencies. Based on our findings, advertisers and agencies would be well advised to examine in detail the role of the brand element in copytesting and to pay due attention to it in the development of copy strategy. Our model can be a useful tool in that process. Second, the identified effects of media planning variables on memory are instrumental in the development of media plans. The effects of media planning variables (serial position and location) on brand memory appeared to vary across the two media contexts that were included in our study. This underlines the importance of accommodating the advertising context in ad location decisions, in which preferred print ad locations should depend on the particular medium choice. Support for advertising strategy decisions, including media class selection, copy execution, editorial fit, scheduling, and timing, require that additional media planning variables and ad characteristics be investigated, which we will do in the near future.[5]

Appendix: MCMC Estimation

The parameters of the model were estimated, using Markov Chain Monte Carlo (MCMC) methods. For each of the distributions of the data and parameters we use a standard parameterization, as in Gelman et al. (1995, pp. 474–477). All prior distributions are standard

[5]This research was sponsored by Verify International, Rotterdam, the Netherlands. The authors thank D.P.G. Claessens of Verify International for gracious allowing us to use the data, and for his support throughout the project. They are grateful for the useful comments of two reviewers, the area editor, and the editor.

conjugate distributions, with the exception of the priors for the parameters of the NBD, which we specify to be normal. In each iteration of the MCMC chain we draw from the full conditional distribution of a parameter q, conditional on the values of the other parameters obtained from the last draw: $\Xi_q \mid \Xi_{/q}, M$. Here $M = \{n, y, t\}$ denotes the data. In developing the MCMC algorithm, we have $\alpha = (\alpha_1', \alpha_2')'$, $\beta = (\eta, \mu')'$, $\Psi^a = (\psi^{a^2} \mid \Sigma^a)$, $\Psi^s = (\psi^{s^2} \mid \Sigma^s)$. We use the following prior distributions:

$$\begin{aligned}
&\alpha_{1j} \sim N(A_0, B_0^{-1}),\ \alpha_{2j} \sim N(A_0, B_0^{-1}),\ \kappa \sim N_3(A_0, B_0^{-1}),\\
&\quad \beta \sim N_4(A_0, B_0^{-1}),\ \rho \sim N_2(A_0, B_0^{-1}),\ \gamma \sim N_2(A_0, B_0^{-1}),\\
&\quad \Psi^{a^{-1}} \sim W(G_0, H_0),\ \Psi^{s^{-1}} \sim G_4(G_0, H_0),\ \xi^{-2} \sim G(G_0, H_0),\\
&\quad \text{with } A_0 = 0,\ B_0 = 10^{-2},\ G_0 = 1,\ H_0 = 10^{-2}.
\end{aligned}$$

We sample the parameters successively from their full conditional posterior distributions, obtained as follows:

1. $\alpha_j \mid n_j, \sim N(A_0, B_0) \times NBD(n_j \mid \alpha_{1j}, \alpha_{2j})$ and $\rho_j \mid n_j, \sim N(A_0, B_0) \times NBD(n_j \mid \rho_j, \alpha_{2j})$. Because the prior and the likelihood are not conjugate, we use a Metropolis-Hastings algorithm to sample from the posterior distribution.

2. $\tilde{D}_{il} \mid y_{il}, \sim \dfrac{N_0^{\infty}(D_{il}, 1 \mid y_{il} = 1)}{N_{-\infty}^{0}(D_{il}, 1 \mid y_{il} = 0)}$, a truncated normal distribution, where $\tilde{D}_{il} = D_{il} + \epsilon_{il}$.

3. $\beta_l^a \sim N((X_l'X_l + \Psi^{a^{-1}})^{-1} (\Psi^{a^{-1}} \beta + X_l'Q_l)\ (X_l'X_l + \Psi^{a^{-1}})^{-1})$, where $X_l = R_l \otimes (1, \gamma_1\xi^{-1})$, with $R_l = (r_{il}) = (1, n_{il}')$ and $Q_l = (Q_{il})$, with $Q_{il}' = (\tilde{D}_{il}, \xi^{-1} (\ln t_{il} - \gamma_0 - \gamma_1 r_{il}\beta_i^s))$.

4. $\beta_i^s \sim N((X_i'X_i + \Psi^{s^{-1}})^{-1} (X_i'W_i), (X_i'X_i + \Psi^{s^{-1}})^{-1})$, where $X_i = R_i \otimes (1, \gamma_1\, \xi^{-1})$, with $R_i = (r_{il}) = (1, n_{il}')$ and, $W_l = (W_{il})$, with $W_{il}' = (\tilde{D}_{il}, \xi^{-1}(\ln t_{il} - \gamma_0 - \gamma_1 r_{il}\beta_l^a))$.

5. $\beta \sim N((L \cdot \Psi^{a^{-1}} + I \cdot \Psi^{s^{-1}} + B_0)^{-1} (\Psi^{a^{-1}} \beta_{\cdot}^a + A_0B_0), (L \cdot \Psi^{a^{-1}} + I \cdot \Psi^{s^{-1}} + B_0)^{-1})$.

6. $\Psi^{a^{-1}} \sim W(L + G_0, (\beta_l^a - \beta)(\beta_l^a - \beta)' + H_0)$.

7. $\Psi^{s^{-1}} \sim W(I + G_0, diag(\beta_l^s\beta_l^{s'}) + H_0)$.

8. $\gamma \sim N((\ln t' \ln t + \xi^{-2})^{-1} (\xi^{-2} A_0 + \tilde{D}' \ln t), (\ln t' \ln t + \xi^{-2})^{-1})$.

9. $\xi^{-2} \sim G(IL/2 + G_0, 1/2\, \Sigma_{i,l} (\ln t_{il} - \gamma_0 - \gamma_1\tilde{D}_{il})^2 + H_0)$.

10. $\kappa \sim N((\beta_0^{a'}\, \beta_0^a + \nu^{-2})^{-1} (\nu^{-2} A_0 + z'\beta_0^a), (\beta_0^{a'}\, \beta_0^a + \nu^{-2})^{-1})$.

11. $\nu^{-2} \sim G(IL/2 + G_0, 1/2\, \Sigma_l (\beta_{0,l}^a - z_l'\kappa)^2 + H_0)$.

In the applications, using starting estimates obtained from maximization of approximate likelihood functions, we ran the MCMC chain for 10,000 iterations, with a burn-in period of 5,000 iterations. Every fifth draw was retained. Stabilization of time series plots of the samples and the stabilization of their quantiles were used to assess convergence. After the burn-in, all chains were stationary. Convergence was checked also by running additional (shorter) chains. We take the posterior median and 95% credible regions as characterizations of the posterior distributions of the parameters. To test the algorithm and investigate identification, we generate two synthetic data sets. Table A1 offers the true parameter values (taken to be of approximately the same magnitude as the estimates in our applications and slightly different for the two data sets; we omit the explanatory variables for the sake of this illustration) and the results. The 95% credible interval covers the true values in all cases.

References

Batra, R., J. G. Myers, D. A. Aaker. 1996. *Advertising Management.* Prentice Hall, Upper Saddle River, NJ.

Bundesen, C. 1990. A theory of visual attention. *Psych. Rev.* **97** 523–547.

Burke, R. R., T. K. Srull. 1988. Competitive interference and consumer memory for advertising. *J. Consumer Res.* **15** 55–68.

Busemeyer, J. R., J. T. Townsend. 1993. Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psych. Rev.* **100** (3) 432–459

Coeffe, C., J. K. O'Regan. 1987. Reducing the influence of nontarget stimuli on saccade accuracy: Predictability and latency effects. *Vision Res.* **27** 227–240.

Cowan, N. 1995. *Attention and Memory: An Integrated Framework.* Oxford University Press, Oxford, U.K.

Finn, A. 1988. Print ad recognition readership scores: An information processing perspective. *J. Marketing Res.* **25** (May) 168–177.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Rubin. 1995. *Bayesian Data Analysis.* Chapman & Hall, London, U.K.

Gilks, W. J., S. Richardson, D. J. Spiegelhalter. 1996. *Markov Chain Monte Carlo in Practice.* Chapman & Hall, London, U.K.

Greene, R. L. 1986. Sources of recency effects in free recall. *Psych. Bull.* **99** 221–228.

Ha, L., B. R. Litman. 1997. Does advertising clutter have diminishing and negative returns? *J. Adv.* **26** 31–42.

Hanssens, D. M., B. A. Weitz. 1980. The effectiveness of industrial print advertisements across product categories. *J. Marketing Res.* **17** (August) 294–306.

Hawley, K. J., W. A. Johnston. 1991. Long-term perceptual memory for briefly exposed words as a function of awareness and attention. *J. Experiment. Psych.: Human Perception Performance* **17** 807–815.

Henderson, J. M. 1992. Visual attention and eye movement control during reading and picture viewing. K. Rayner, ed. *Eye Movements and Visual Cognition.* Springer Verlag, New York, 260–283.

Hintzman, D. L. 1988. Judgements of frequency and recognition memory in a multiple-trace memory model. *Psych. Rev.* **95** (4) 526–531.

Jacoby, L. L., M. Dallas. 1981. On the relationship between autobiographical memory and perceptual learning. *J. Experiment Psych.: General* **110** 306–340.

Janiszewski, C. 1990a. The influence of nonattended material on the processing of advertising claims. *J. Marketing Res.* **27** 263–278.

——. 1990b. The influence of print advertisement organization on affect toward the brand name. *J. Consumer Res.* **17** 53–65.

——. 1998. The influence of display characteristics on visual exploratory search behavior. *J. Consumer Res.* **25** 290–301.

——, B. Bickart. 1994. Managing attention. *Adv. Consumer Res.* **21** 329.

Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lütkepohl, T. C. Lee. 1985. *The Theory and Practice of Econometrics.* Wiley, New York.

Just, M. A., P. A. Carpenter. 1988. Reading and spatial cognition: Reflections from eye fixations. G. Luer, U. Lass, J. Shallo-Hoffmann, eds. *Eye Movement Research: Physiological and Psychological Aspects.* C. J. Hogrefe, Toronto, Canada, 193–213.

Kahneman, D. 1973. *Attention and Effort.* Prentice Hall, Englewood Cliffs, NJ.

Karslake, J. S. 1940. The Purdue eye-camera: A practical apparatus for studying the attention-value of advertisements. *J. Appl. Psych.* **24** 417–440.

Krishnan, H. S., D. Chakravarti. 1993. Varieties of brand memory induced by advertising: Determinants, measures, and relationships. D. A. Aaker, A. L. Biel, eds. *Brand Equity and Advertising: Advertising's Role in Building Strong Brands.* Lawrence Erlbaum, Hillsdale, NJ 213–231.

——, ——. 1999. Memory measures for pretesting advertisements: An integrative conceptual framework and a diagnostic template. *J. Consumer Psych.* **8** 1–37.

——, S. Shapiro. 1996. Comparing implicit and explicit memory for brand names from advertisements. *J. Experiment. Psych.: Appl.* **2** 147–163.

Kruschke, J. K. 1996. Base rates in category learning. *J. Experiment. Psych.: Learn. Memory Cognition* **22** (1) 3–26.

LaBerge, D. 1995. *Attentional Processing: The Brain's Art of Mindfulness.* Harvard University Press, Cambridge, MA.

Loftus, G. R. 1972. Eye fixations and recognition memory for pictures. *Cogn. Psych.* **3** 525–551.

——, N. H. Mackworth. 1978. Cognitive determinants of fixation location during picture viewing. *J. Experiment. Psych.: Human Perception Performance* **4** 565–572.

McConkie, G. W. 1983. Eye movements and perception during reading. K. Rayner, ed. *Eye Movements in Reading: Perceptual and Language Processes.* Academic Press, New York, 65–96.

——, P. W. Kerr, B. P. Dyre. 1994. What Are "normal" eye movements during reading: Toward a mathematical description. J. Ygge, G. Lennerstrand, eds. *Eye Movements in Reading.* Elsevier Science Inc, Tarrytown, NY. 315–327.

Ober, J. 1994. Infra-red reflection technique. J. Ygge, G. Lennerstrand, eds. *Eye Movements in Reading.* Elsevier Science Inc, Tarrytown, NY, 9–19.

Olney, T. J., M. B. Holbrook, R. Batra. 1991. Consumer responses to advertising: The effects of ad content, emotions, and attitude toward the ad on viewing time. *J. Consumer Res.* **17** 440–451.

Pieters, R., E. Rosbergen, M. Wedel. 1999. Visual attention to repeated print advertising: A test of scanpath theory. *J. Marketing Res.* **36** (November).

——, L. Warlop. 1999. Visual attention during brand choice: The impact of time pressure and task motivation. *Internat. J. Res. Marketing* **16** 1–16.

Raaijmakers, J. G. W., R. M. Shiffrin. 1992. Models for recall and recognition. *Ann. Rev. Psych.* **43** 205–234.

Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psych. Bull.* **124** 372–422.

Richardson-Klavehn, A., R. A. Bjork. 1988. Measures of Memory. *Ann. Rev. Psych.* **39** 475–543.

Rosbergen, E., R. Pieters, M. Wedel. 1997. Visual attention to advertising: A segment-level analysis. *J. Consumer Res.* **24** 305–314.

Rossiter, J. R. 1981. Predicting starch scores. *J. Adv. Res.* **21** (October) 63–68.

——, L. Percy. 1997. *Advertising Communications and Promotion Management*, 2nd ed. McGraw-Hill, New York.

Sandage, C. H. 1945. *Advertising: Theory and Practice.* Richard D. Irwin, Inc., Chicago, IL.

Sergent, J., Y. Takane. 1987. Structures in two-choice reaction time data. *J. Experiment. Psych.* **13** 300–315.

Sperling, G., E. Weichselgartner. 1995. Episodic theory of the dynamics of spatial attention. *Psych. Rev.* **102** 503–532.

Suppes, P. 1994. Stochastic models of reading. J. Ygge, G. Lennerstrand, eds. *Eye Movements in Reading.* Elsevier Science Inc, Tarrytown, NY, 349–364.

Viviani, P. 1990. Eye movements in visual search: Cognitive, perceptual and motor control aspects. E. Kowler, ed. *Eye Movements and Their Role in Visual and Cognitive Processes.* Elsevier Science Publishers, Amsterdam, The Netherlands, 353–393.

Wyer, R. S., T. K. Srull. 1986. Human cognition in its social context. *Psych. Rev.* **93** 322–359.

This paper was received October 5, 1999, and has been with the authors 3 months for 2 revisions; accepted by Greg Allenby.