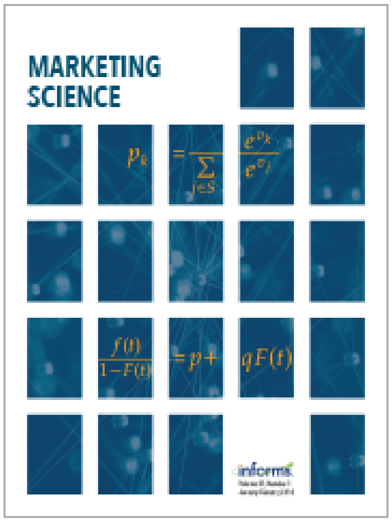



Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

# A Neurocognitive Model of Advertisement Content and Brand Name Recall

Antonio G. Chessa, Jaap M. J. Murre,

Marketing Science
Vol. 26, No. 1, January–February 2007, pp. 130–141
issn 0732-2399 | eissn 1526-548X | 07 | 2601 | 0130

doi 10.1287/mksc.1060.0212


# A Neurocognitive Model of Advertisement Content and Brand Name Recall

Antonio G. Chessa
Statistics Netherlands (CBS), P.O. Box 4000, 2270 JM, Voorburg, The Netherlands, antoniogchessa@yahoo.com

Jaap M. J. Murre
Department of Psychonomy, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands, jaap@murre.com

We introduce a new (point process) model of learning and forgetting, inspired by the structures of the brain, that we apply to model long-term memory for advertising and brand name recall. Recall-probability functions derived from the model are tested with classic data by Zielske [Zielske, H. A. 1959. The remembering and forgetting of advertising. *J. Marketing* **23** 239–243], as well as advertisement content and brand name recall data of a Dutch study that tracked over 40 campaigns of TV commercials. Data fits and cross-validation results indicate that the recall functions serve as a good first approximation for aggregate behavior. The shapes of optimal GRP schedules, which are obtained by maximizing a recall measure, are strongly related to the model parameters and corresponding memory processes. Comparisons with existing models in the literature indicate that a neurobiologically motivated model may give a more realistic description of memory for advertisements.

*Key words*: advertising; memory; impact; scheduling; bursting; dripping; massed and spaced learning; point processes
*History*: This paper was received May 28, 2003, and was with the authors 27 months for 3 revisions processed by Joel Huber.

## Introduction

Should advertisements be concentrated in time, spread out evenly, or distributed according to a mixture of these two strategies (e.g., Merzereau and Battais 2000)? A similar question is studied in the field of memory psychology: How should we learn new material such as foreign vocabulary words, according to a *massed* (bursting) or *spaced* (dripping) schedule? Most studies conclude that spaced learning is superior; others report that the optimal performance is obtained by some combination of massed and spaced learning (e.g., Baddeley and Longman 1978, Bahrick et al. 1993, Glenberg 1976, Glenberg and Lehmann 1980, Rumelhart 1967).

Memory performance for advertising material shows results similar to those obtained in memory psychology. Burtt and Dobell (1925) were the first to study the recall of advertisements; their results indicated a form of dependence between the interpresentation lag and retention lag. The classic study by Zielske (1959) ranged over a period of one year, during which the recall of printed advertisements was measured during 13 repeated mailings that were sent to subjects either every week or every four weeks. The spaced, four-week schedule proved to produce a significantly higher memory performance. Although the superiority of the spaced schedule is evident from Zielske's study, these findings are not sufficient to assist decision makers in selecting optimal advertising schedules. Is every four weeks the optimal schedule, or is the optimum located between one and four weeks, or at a larger spacing? Is this universally true, or does it depend on the brand and campaign?

The similarity between the results on this so-called spacing effect in the psychological and advertising literature, and the difficulty with extrapolating these empirical results to real-world situations, has motivated us to develop a mathematical memory model, the *memory chain model*. This model can be seen as an abstraction and extension of a neurological model of long-term memory by our group (Murre 1996; Murre et al. 2001; Meeter and Murre 2004, 2005; Murre et al. 2001; Talamini et al. 2005). In this paper, we will use the model to derive expressions for advertisement content and brand name recall. These recall functions will be fitted and validated on about 80 data sets. After the validation, we will use the fitted recall measures to optimize advertisement scheduling. Finally, we will study to what extent schedule type and short-term and long-term memory parameters correlate. We emphasize here that we will focus only on improving advertisement scheduling, not on improving the advertisement message.

In the next section, we will describe the most relevant details of the memory chain model. Our approach has a number of novel aspects: (1) the model formalizes neurobiological processes based on recent findings, and (2) we use a single model of memory to derive different expressions for learning and forgetting of both advertisement content and brand name recall. The memory chain model has already been applied as such to hundreds of memory retention data sets (e.g., Chessa and Murre 2004, Janssen et al. 2005).

There are a few models in the psychological literature that have some similarity to the memory chain model (Atkinson and Shiffrin 1968, Murdock 1974). However, these models only describe situations with a single learning trial and subsequent forgetting, in which the short-term to long-term memory processes are rather restrictive. While the existing models are limited to one or two memory stores, the memory chain model can handle any number of stores. Two other types of modeling approaches are based on: (1) Markov chains (Murdock 1974), where the states represent the level of mastery of some item memorized; and (2) regression procedures for fitting functions to retention data (Rubin and Wenzel 1996). In the latter approach, the functions used are not derived from an underlying formal theory of memory, as is the case with the memory chain model, but are simply selected and fitted. Markov models only formalize the effect of learning at successive trials, but do not model forgetting between intermediate trials, which the memory chain model does take into account.

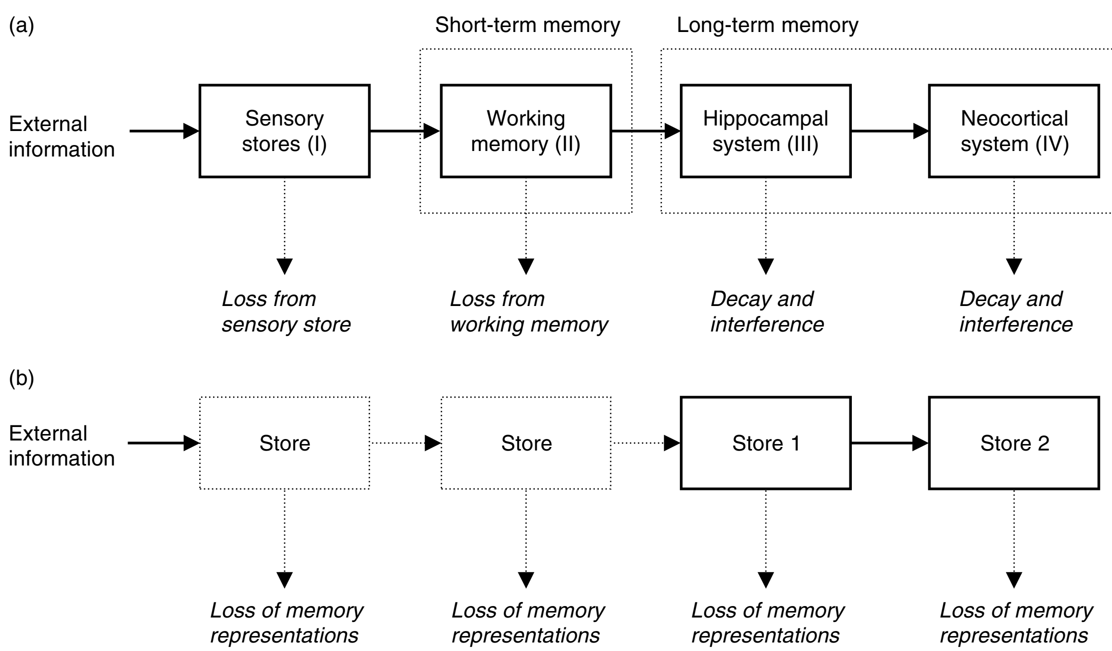
The data on which the recall functions for advertisement content and brand name will be tested and validated are presented in the third and fourth sections. The results of the data fits and optimization will be discussed in the discussion section, where we will also compare the results with a model presented in Naik et al. (1998).

## The Memory Chain Model

The memory model is built upon two main concepts. During exposure (e.g., to an advertisement), a memory is encoded as a number of *representations*, each of which captures characteristics of the memorized item. Representations may be activated over time by a memory cue, for instance, by a product category when a brand name has to be recalled. We model the process of activated memory representations in time as a point process (Daley and Vere-Jones 1988, Diggle 1983, Stoyan et al. 1987). We will consider a *Poisson point process*; that is, numbers of representations activated in time intervals have a Poisson distribution, and numbers in disjoint intervals are independent. Point processes have been used for modeling neural spike trains (Abeles 1991).

The second concept is motivated by neurobiological evidence that memories may be stored in one or more parts of the brain, which we will denote in abstract terms by *memory stores*. Two processes characterize each store: a process of memory decline or loss, and an induction process, which transfers representations from one store to another in a feed-forward fashion (Figure 1). A memory can be strengthened in a second store, for example, through induction by activation. This is an induction process where, for example, a rapidly decaying neural group activates neurons in a less rapidly decaying part of the brain. An exam-

**Figure 1** (a) Storage Systems for a Memory at Different Time Scales, with Feed-Forward Induction Between and Decline Within Stores. (b) Abstract Representation Used in the Memory Chain Model



ple of the latter is the dorsolateral prefrontal cortex, which is assumed to hold a working memory representation for a few seconds to minutes (Goldman-Rakic 1992, 1995). From a neurobiological point of view, these processes are part of a cascade of induction and decline processes that take place at different time scales (McGaugh 2000).

We will use models with two stores, at most, in the fits to advertising data in order to limit the number of parameters. Because all the data fall in time ranges in the order of weeks and months, the two stores of the model will most likely reflect neural processes in the hippocampus (Store 1) and the neocortex (Store 2). There are no data at time lags shorter than a week, so that the short-term stores in the upper panel of Figure 1 are in fact covered by Store 1 in our model fits. It should therefore be noticed that there is not necessarily a one-to-one correspondence between the neural stores (Figure 1(a)) and the stores of the memory chain model (Figure 1(b)).

### A Single Learning Trial

The memory chain model formalizes memory and retention as a process that consists of four stages, which are described below.

**Encoding.** During stimulus exposure, the number of memory representations eventually reaches a mean value, or *intensity*, $\mu_1$. Encoded memory representations are stored in the first memory store (i.e., Store 1). Representations are not labeled, so the model formalizes how much is stored. Memory representations could be seen as critical features or relevant details of an advertisement.

**Storage.** After encoding, a decline process is activated, for example, by spontaneous decay, by the overwriting of other learned items, by neural noise, or by other factors. An encoded memory representation is still available at time $t$ since the end of a learning trial with probability denoted by $\tilde{r}_1(t)$, which will be called the *decline function* of Store 1. Memory in Store 1 is assumed to be a *nonhomogeneous Poisson process* with *intensity function* $r_1(t) = \mu_1 \tilde{r}_1(t)$. We will assume an exponential decline function, so that

$$r_1(t) = \mu_1 \exp(-a_1 t), \tag{1}$$

where $a_1 > 0$. Data obtained from laboratory experiments that intend to measure short-term memory decline through the classical Brown-Peterson learning and distraction task support an exponential decline (Peterson and Peterson 1959, Murdock 1961). Evidence at the neural level is offered by the long-term potentiation or synaptic growth and decline data in the hippocampus of young and old rats (Barnes and McNaughton 1980).

Memory representations may be induced from Store 1 to Store 2, where they form a new point process. This point process arises from a time-series of *inductions* (which in some experiments can be thought of as rehearsals), of which the intensity is proportional to the intensity in the first store, that is, $\mu_2 \mu_1 \tilde{r}_1(t)$, where $\mu_2$ is a (time-independent) induction rate. An induction that occurs at time $\tau$ induces a memory representation at time $t \geq \tau$ according to probability density function $a_2 \exp(-a_2(t-\tau))$ in Store 2, which we write as $f$. This function specifies the decline rate in the second store, which we assume to be smaller than in the first store ($a_2 < a_1$). The intensity function $r_2$ of the second store is a convolution of $\mu_2 \mu_1 \tilde{r}_1(t)$ and $f$, which yields the expression:[1]

$$r_2(t) = \frac{\mu_1 \tilde{\mu}_2}{a_1 - a_2}(\exp(-a_2 t) - \exp(-a_1 t)), \tag{2}$$

where $\tilde{\mu}_2 = \mu_2 a_2$, for which we will continue to write $\mu_2$.

**Retrieval.** Memory stores are searched for evidence at a recall test. Stores can be searched either simultaneously or sequentially; it can be proved that both strategies lead to identical expressions for recall probability. The effectiveness of the search process and of the retrieval cue(s) used for the search is expressed by a single parameter $q$, which can be interpreted as the probability that a memory representation will be cued. Cueing has the following effect on the total intensity $r_1 + r_2$ over the two stores:[2]

$$r(t) = (r_1(t) + r_2(t))q. \tag{3}$$

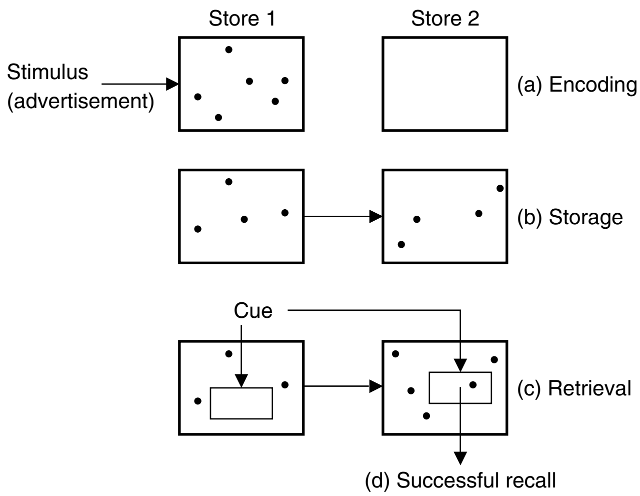
This model characteristic does not necessarily imply that cue effectiveness is constant over time. The retention functions that follow from our memory model are invariant under exponentially varying cue effectiveness. This assumption does not lead to additional degrees of freedom in the model, because cue effectiveness can be subsumed under the exponential memory decline functions of the stores. We will combine $q$ with the encoding parameter $\mu_1$, for which we will continue to write $\mu_1$, because the effect of cueing is not varied systematically in the studies reported here.

**Recall.** Our model includes a fourth stage, which can be considered a decision process that relates retrieval to recall. The number of retrieved representations is compared with a threshold $b$. Unless stated otherwise, we assume that successful recall takes place if a store exists with at least one cued representation. In some cases, subjects require more than one representation for recall. From laboratory exper-

[1] For more details, see http://www.neuromod.org/staff/murre/mcmforgettingdraft.doc.

[2] The formalisms of memory decline and cueing, as shown in (1) and (3), are in fact *independent thinnings* of Poisson processes with intensities $\mu_1$ and $r_1 + r_2$, respectively, which again result in Poisson processes with thinned intensities (Stoyan et al. 1987).

**Figure 2** Illustration of Memory Processes in a Two-Store Model




*Notes.* (a) Memory representations (points) are generated in Store 1. (b) While memory declines in Store 1, representations are generated in a second store by induction. (c) A cue searches areas in both stores and finds a memory representation in Store 2. (d) A correct response will be given at a recall threshold of 1.

iments, we obtained a strong suggestion that a new stimulus has a higher threshold and thus requires more representations than a well-known stimulus does. From the properties of Poisson processes, it follows that the *recall-probability function* at retention lag $t$ since stimulus exposure takes the form

$$p(t) = 1 - \sum_{n=0}^{b-1} \frac{(r(t))^n}{n!} \exp(-r(t)). \tag{4}$$

For the special case $b = 1$, we thus have $p(t) = 1 - \exp(-r(t))$. The concepts of memory encoding, storage, retrieval, and recall are illustrated in Figure 2.

**Multiple Learning Trials**

We assume that different learning trials give rise to independent Poisson processes, where memory representations generated by different learning trials follow the same storage and retrieval processes as in the single-trial case. The only opportunity for interaction of successive learning trials is at encoding. From a neurobiological point of view, it is rather uncommon that encoding increases linearly as learning continues. Rather, it will saturate, and therefore be limited by some upper bound. For example, the number of firing neural groups cannot become infinitely large.

We will, therefore, assume that the initial encoding may saturate at some finite value $r_{\max} \geq \mu_1$. Let $v$ denote the learning rate per unit of learning time and $l$ denote the duration of the first learning trial. We then describe the initial encoding during the learning trial as follows:[3]

$$\mu_1(l) = r_{\max}(1 - e^{-vl/r_{\max}}). \tag{5}$$

When $\mu_1$ is close to saturation, learning no longer has a significant effect, which is what we expect in memory psychology with prolonged massed learning.

The encoding intensity $\mu^{(L)}$ regarding the individual contribution of subsequent trials $L \geq 2$ is similar to (5). Let $t_i$, $i = 1, 2, \ldots,$ denote the presentation times of a series of learning trials. The encoding $\mu^{(L)}(l)$ of trial $L$ with duration $l$ is given by:

$$\mu^{(L)}(l) = \left(r_{\max} - \sum_{i<L} \mu^{(i)} \tilde{r}_1(t_L - t_i)\right)(1 - e^{-vl/r_{\max}}). \tag{6}$$

Notice that this expression is equal to $\mu_1(l)$, given by (5), multiplied by the relative intensity that can still be accommodated in Store 1 at trial $L$.

The recall probability $p(t)$ at a time lag $t$ since the last trial in a series of $L$ trials has the following form for recall threshold $b = 1$:

$$p(t) = 1 - \exp\left\{-\sum_{i=1}^{L} \mu^{(i)} (\tilde{r}_1(t_L - t_i + t) + \tilde{r}_2(t_L - t_i + t))\right\}, \tag{7}$$

where $\tilde{r}_2 = r_2/\mu_1$, with $r_2$ equal to (2). That is, $\tilde{r}_2$ denotes the decline function of Store 2. Other values of $b$ lead to expressions similar to (4).

**Existing Memories**

It is easy to imagine situations where a memory has already been formed as a consequence of previous exposures to the same item, which may still exist during a new campaign. In our model, we account for such memory representations by assuming a memory that does not decline within the tracking period. We model this as a homogeneous Poisson process with (*base rate*) intensity $\mu_0$, which is assumed to be independent of the point processes induced by new campaigns. The implication of this additional process for recall probability is that the intensity $\mu_0$ is summed with the intensity functions for new campaigns in the exponent of (7). A list of the parameters, functions, and variables in our model is given in Table 1.

We will now present two data sets to which we will apply our model. Recall-probability function (7) will be used in both applications as a first approximation to aggregate recall behavior concerning advertisement content and brand names. This means that we make

[3] We use the term learning in a broad psychological sense, and it can be interchanged with encoding. Consumers could watch an advertisement passively (low $v$) or with a lot of attention (high $v$).

**Table 1 Parameters, Functions, and Variables in the Memory Chain Model**

| | Theoretical meaning | Meaning/effect in applied context |
|---|---|---|
| Parameters | | |
| $\mu_0$ | Intensity of homogeneous (base rate) Poisson process | Expected number of memory representations still left from previous campaigns |
| $v$ | Learning rate per unit of learning time in time-continuous setting | Expected number of memory representations formed per GRP during encoding of ads |
| $\mu_1$ | Initial encoding (intensity after first trial) | Expected number of memory representations at the end of the first ad exposure or GRP pulse |
| $r_{\max}$ | Upper bound on the intensity of Store 1 ($r_{\max} \geq \mu_1$) | Finite $r_{\max}$ leads to decreasing encoding over ad exposures and to equal contributions otherwise |
| $\mu_2$ | Induction rate from Store 1 to Store 2 | Constant transfer rate of memory representations of ads from Store 1 to Store 2 |
| $a_1, a_2$ | Decline rates in Stores 1 and 2, respectively | Relative loss of memory representations in Stores 1 and 2 |
| $b$ | Recall threshold | Minimum number of retrieved representations, e.g., critical features of ads, needed for successful recall |
| Functions | | |
| $r(t)$ | Intensity function summed over stores | Specifies expected number of memory representations taken over both stores as a function of test lag $t$ |
| $r_1(t), r_2(t)$ | Intensity functions of Stores 1 and 2 | Same as $r(t)$, but for Stores 1 and 2 separately |
| $\tilde{r}_1(t), \tilde{r}_2(t)$ | Decline functions of Stores 1 and 2 | Express changes in the encoded number of memory representations for Stores 1 and 2 at test lags $t$ |
| $\mu^{(L)}(l)$ | Encoding for trial $L$ with presentation time $l$ | Expected number of memory representations for ad exposure $L$ with $l$ GRPs |
| $p(t)$ | Recall-probability function | Probability of successfully recalling the content or the name of a brand in a commercial at test lag $t$ |
| | | Meaning in applied context |
| Variables | | |
| $t$ | Retention lag or interval | Time since last ad exposure |
| $l$ | Learning or exposure time | GRPs in SPOT data |
| $t_i$ | Time at which learning trial $i$ is presented | Moment at which an ad is exposed for the $i$th time |
| $L$ | Number of learning trials | Number of advertisements or GRP pulses in campaign |

the simplifying assumption that memory processes are the same for different subjects. We make this choice because we propose a new model in this paper, and we want to find out to what extent this model is already able to fit the data with a limited number of parameters.

## The Zielske Study

### The Data

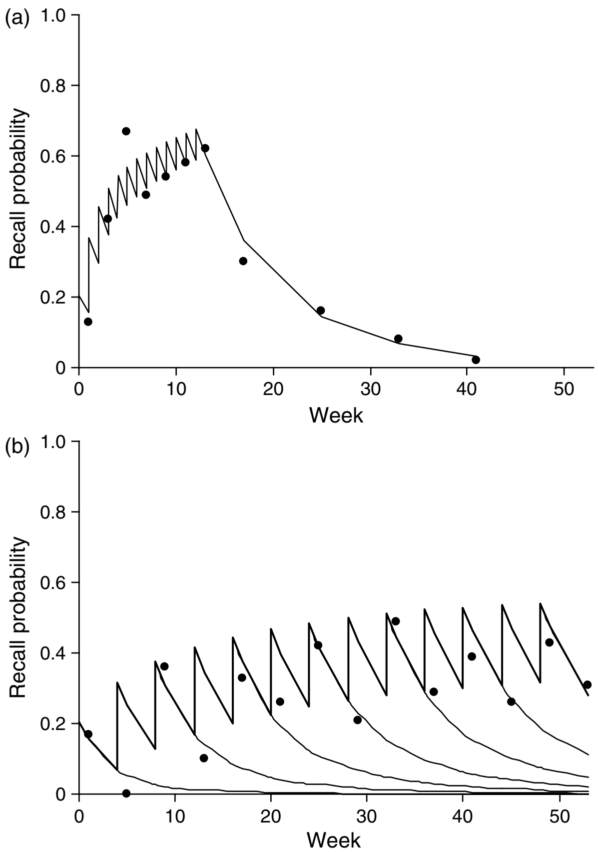
In the literature known to us, the only extensive study about the recall of printed advertising material for repeated advertisements is the classic study by Zielske (1959). Housewives were repeatedly sent the same printed advertisement, according to two schedules. One group received the advertisement once a week for 13 successive weeks, while the second group received the advertisement 13 times with a uniform spacing of four weeks. To obtain data about the recall of the advertisement in time, both groups were subdivided into smaller groups that were interviewed in specific weeks. The interviews were organized in such a way that data about learning and forgetting between and after the exposures were obtained. Simon (1979) analyzed the data and concluded that the one-week schedule reaches a higher peak in recall, but that after about 17 weeks, recall improves for the four-week schedule and stays at a higher level after the end of the campaign. The average recall percentage calculated over a period of one year is higher for the four-week schedule. This implies that the spaced schedule would be preferred on two criteria, namely, (1) mid-term and long-term recall percentages after the end of an advertising campaign, and (2) the average recall percentage calculated over a period of one year.

### The Fitted-Recall Functions

We fitted recall-probability function (7), which holds for recall threshold $b = 1$, and recall-probability functions with higher recall thresholds, which have the form of (4). A function was fitted simultaneously to both schedules. This means that we used the same parameter values in the fits to both schedules, because differences in recall for two different schedules should be explained by one underlying memory mechanism, and should merely represent the effect of intermediate forgetting between two successive advertisements.

The parameters were estimated by minimizing a chi-square statistic, which is the sum of conventional chi-square statistics computed for every advertising

**Figure 3** Fits of a Two-Store Model with Saturated Learning to Zielske's Recall Data (Dots) of a Printed Advertisement




*Notes.* (a) Fit to the one-week advertising recall data: The first seven data points refer to 1, 3, 5, 7, 9, 11, and 13 advertisements received. The last four points denote forgetting after the 13th and final advertisement. (b) Fit to the four-week advertising recall data: The data points refer to 1, 3, 5, 7, 9, 11, and 13 advertisements, respectively, with each number of advertisements occurring twice in succession. The thin curves denote the course of forgetting after these numbers of advertisements.

week. The degrees of freedom are equal to the difference between the number of advertising weeks (i.e., data points) and the number of parameters estimated (Wickens 1982). From the minimum chi-square statistic we computed the size $\alpha$ for each fitted function at which a null hypothesis is not rejected. This approach was also carried out for the SPOT data in the next section.

Figure 3 shows the fit of a two-store model with saturated learning over advertisement exposures and a recall threshold $b = 2$. This was the only version of our model to fit the data for test sizes greater than 0.05 ($R^2 = 0.96$). The fits show the typical *saw-tooth* behavior, with recall peaks at each exposure and decline between successive exposures. Respondents were asked three different questions about the advertisement in order to test recall. This is reflected by the value of the recall threshold, which implies that they had to recall more than one feature of the advertisement.

The expected recall percentage over a one-year period can be calculated exactly for the two schedules simply by summing the expected recall over the weeks. On the basis of this criterion, the four-week schedule is one-and-a-half times more effective. An advantage of using a model is that other schedules can be defined and compared. We analyzed the expected yearly performance for schedules with both shorter and longer intervals, but the four-week schedule gave the highest one-year recall. Because of the saturation of advertisement memory, spaced advertising gives a better performance than massed advertising.

## The SPOT Tracking Study

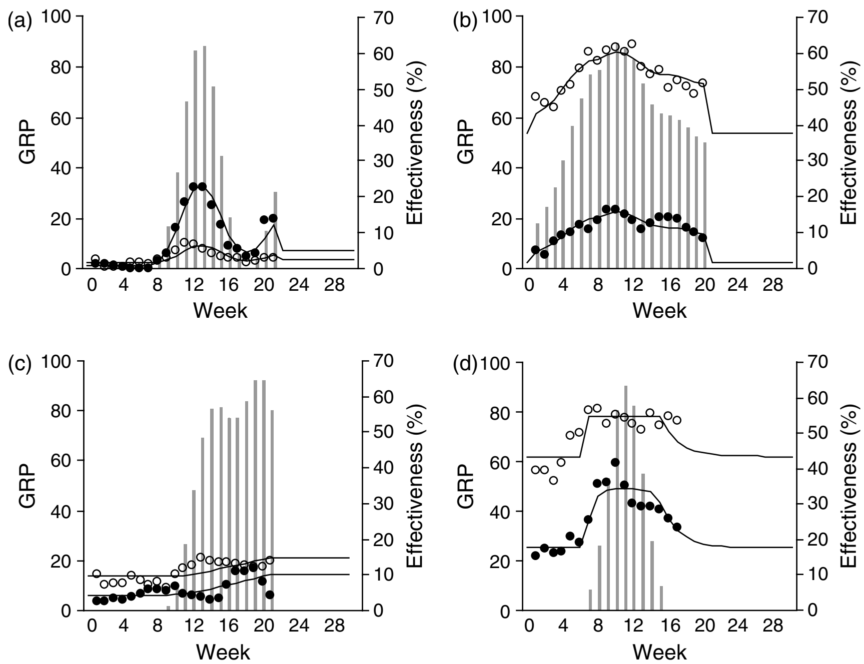
### The Data

In 1997–1998, the Foundation for Promotion and Optimization of Television Advertising in the Netherlands (SPOT 1998) carried out an extensive study on the effectiveness of television advertising in the Netherlands. Product categories were considered with a large advertising budget, using television as the primary medium. A total of 42 brands within the categories "detergents" and "food" were tracked in campaigns that were scheduled in the period May 1997 until December 1997. (We do not mention brand names for reasons of confidentiality.) Data were collected through 50 phone calls per week for every brand. Only those respondents who did most of the shopping in their family and were older than 18 were interviewed. The results obtained were adjusted on the basis of age, education, residence, and type of household. Data about different advertising effectiveness measures were collected, of which we will analyze the following:

- *Impact*: This is proven recall about the contents of an advertisement. Respondents were asked, in an open-ended question, what they recalled about the TV commercial of a brand. The answers were coded according to a standard that was developed for this study and converted to a score between 0% and 100%;
- *Brand name recall*: This is the spontaneous recall of one or more brand names that belong to a specific product group. The interviewer named the product group, after which the respondent was asked to recall as many corresponding brand names as came to mind.

The data for these two effectiveness measures are presented in Figure 4 for a small selection of brands. Buying intention was measured as well in the SPOT study, but will be analyzed elsewhere, as this measure requires the integration of our model with a decision-analytic framework for comparing and ranking brands, which is beyond the scope of the present study.

**Figure 4** Impact (Solid Circles) and Brand Name Recall (Open Circles) as a Function of GRPs (Gray Bars) for Four Brands (a: Meal, b: Detergent, c: Meal Sauce, d: Candy)




*Notes.* The model fits are shown as lines (thick line for impact, thin line for brand name recall): (a) a two-store model without saturation, rapid forgetting in Store 1, slow forgetting in Store 2, and a recall threshold equal to 1; (b) strong forgetting, no saturation; (c) no forgetting, no saturation; (d) strong saturation with slow forgetting.

## The Fitted-Recall Functions

In Zielske's study, every individual was exposed to an advertisement once in an advertising week. In the SPOT study, individuals may have been exposed more than once, or never, to a commercial, which gives rise to a mean exposure frequency or GRP per week. We will approximate aggregate recall behavior by substituting learning time $l$ in (5) and (6) by GRP for every advertising week or learning trial $L$.

Because existing commercials were also tracked, the base rate $\mu_0$ for existing memories was included in the fits to both impact and brand name recall data. The remaining model assumptions with respect to impact are the same as for the model fitted to the Zielske data, that is, we fitted one-store and two-store models with the recall threshold as a free parameter. We fitted the same functions used for impact to brand name recall, subject to a number of restrictions. In contrast with the contents of the TV commercials, we assumed that all names of the 42 brands could be recalled easily, because they mostly consist of one word. The recall threshold $b$ was therefore set equal to one in each of the 42 fits. The base rate $\mu_0$ for existing memories may be different with respect to the base rate for impact, because a new commercial may be launched for a well-known brand, and because other media may also contribute to brand name recall. We also allowed the learning rate $v$ to be different between brand names and other details of a commercial. The familiarity already gained with a brand name through previous advertisements may decrease the contribution of TV commercials on its encoding. The remaining parameters were fixed at the same value obtained in the model fits to impact data. We thus have only two free parameters in the fits for brand name recall.

Fits to the SPOT data on impact and brand name recall are shown in Figure 4. On a total of 84 chi-square goodness-of-fit tests, three fits were rejected, at $\alpha = 0.05$, for both impact and brand name recall, which amounts to about 93% of nonrejections. We think that the rejections are caused by two factors: (1) nonuniform competition effects from other brands; (2) recall is not constant in the weeks before the TV-advertising campaign for certain brands, which indicates effects of possible advertising through other media. These effects are represented by the base rate parameter $\mu_0$, which we assumed to be constant in time.

A two-store model gave the best fit to the impact data for 11 brands. In Figure 4(a), the impact and brand name recall fits show that recall declines after the end of the campaign to values that are greater than the precampaign base rates, indicating memory consolidation to a second store. Of the remaining data sets, 28 were fitted with a one-store model, and three required only the stationary base rate process, which means that these three campaigns had no effect on impact.

A two-store model gave the best fit to brand name recall for seven campaigns, and 20 gave the best fit with a one-store model. Fifteen brands only required the stationary base rate process. For these brands, the corresponding advertising campaigns did not have any effect on brand name recall. The results thus show numerous examples of TV commercials that improve impact but not brand name recall.

The question that we will try to answer in the next subsection is whether the memory processes that underlie the fitted-recall functions imply GRP schedules that improve the average impact and brand name recall over the tracking periods. For example, Figure 4(d) shows a brand with strong saturation, which is evidenced by the flat part of the impact and brand name recall functions during repeated advertising. Should one choose a schedule with an increased number of exposures, as is shown in Figure 4(d) during the first half of the campaign period; or a uniformly spaced schedule over a longer advertising period, given the decreasing contributions of repeated exposures to recall? Before answering such questions, we will subject the recall functions to a cross-validation study.

To assess the predictive performance of recall functions, we perform a cross-validation study, where a part of the data is used to estimate model parameters and the remaining data is predicted by using the functions fitted to the restricted data sets. We left out the data for the last five campaign weeks for every brand. Table 2 shows the values of three statistics for the recall functions fitted to the complete impact data sets and to the restricted data sets used for the cross validation. We also listed the results for the model proposed by Naik et al. (1998). Table 2 shows average values over all brands in each of four product classes.

We included $R^2$ because it is widely used in model fits. However, we would like to emphasize that it does not always give clear indications about the quality of a fit. A low $R^2$ value may correspond to test sizes $\alpha > 0.05$. This depends on factors such as the sample size per data point and the number of model parameters used in a fit. For example, a data set with small variation may be fitted well by a constant function, as additional parameters could decrease the test size $\alpha$ in a composite test. The result will then be a fit with a low $R^2$ value, which will increase with additional parameters, causing an opposite effect compared to $\alpha$. Such fits also influenced the results in Table 2, as a comparison between the values for maximum $\alpha$ (which corresponds with the minimized chi-square statistic) and the number of rejected functions on the one hand, and the $R^2$ values on the other hand, illustrates.

The memory chain model gives better fits than the model by Naik et al. (1998). We will analyze and compare the two models in the discussion section. Table 2 also indicates that the model used for the cross validation gives results that are in close agreement with the model fitted to the complete data sets. The parameter estimates turned out to be stable when comparing the estimates for the restricted data and the full data. In more than 70% of the brands, the relative change in the parameter values remained within 10%. The fitted functions were hardly affected in the majority of the remaining cases, because the estimated parameter values were very small.

### Optimization of Advertisement Scheduling

We used the recall-probability functions that were fitted to impact and brand name recall to investigate whether more effective campaigns exist with a different GRP scheduling, under the same total amount of GRP and campaign period. Different criteria can be chosen for optimization: impact and brand name recall in a specific target week (e.g., when advertising a movie premiere), or the average impact and brand name recall over a time interval (stimulate sales consistently). Here we focus on the latter.

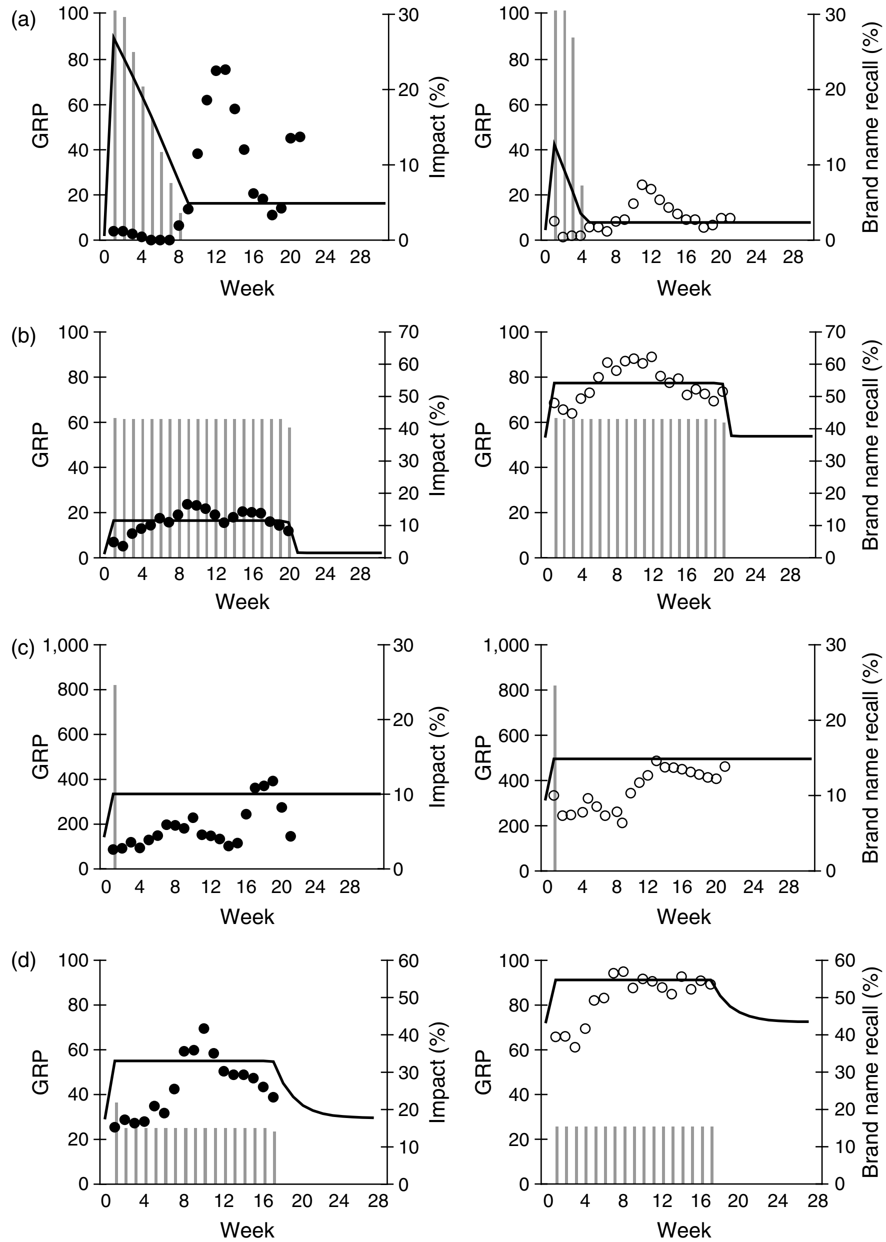
An increase of the average expected impact by more than 2% was found for six brands, and by more than 5% for two brands. Regarding brand name recall, two brands improve by 4%–5%, while the remaining brands improve up to 1%. A brand for which the average impact improves significantly under the optimal GRP distribution is a brand that is fitted with a one-store model with saturation and rapid forgetting (Figure 5(d)). For this brand, the average expected

**Table 2** **Values of Three Statistics for Three Models Fitted to the Impact Data (SPOT): Average Maximum Test Size $\alpha$ for Minimized Chi-Square Statistics, the Number of Rejected Functions ($\alpha = 0.05$), and Average $R^2$ (for Brands Grouped into Four Product Classes)**

| | MCM (cross-validation) | | | MCM (complete fit) | | | NMS (complete fit) | | |
|---|---|---|---|---|---|---|---|---|---|
| Product class | Alpha | Rejected | *R*-square | Alpha | Rejected | *R*-square | Alpha | Rejected | *R*-square |
| Detergents (18) | 0.56 | 1 | 0.44 | 0.55 | 0 | 0.50 | 0.46 | 4 | 0.49 |
| Ice and coffee (8) | 0.75 | 0 | 0.79 | 0.63 | 0 | 0.79 | 0.67 | 1 | 0.76 |
| Meal products (11) | 0.64 | 2 | 0.48 | 0.53 | 3 | 0.49 | 0.32 | 4 | 0.32 |
| Candy (5) | 0.59 | 0 | 0.67 | 0.59 | 0 | 0.82 | 0.50 | 0 | 0.81 |

*Notes*. The number of brands per product class is given between parentheses in the first column. MCM = memory chain model; NMS = model of Naik et al. (1998).

**Figure 5 GRP Distributions (Gray Bars) for Maximized Average Expected Impact (Solid Line, Left) and Brand Name Recall (Solid Line, Right) over the Campaign Period**




*Notes.* Data are for the same fitted functions as in Figure 4. The original impact and brand name recall data (dots) are also shown.

impact increases from 26.3% to 33%, while brand name recall increases from 50.1% to 54.8% under the optimal schedule. Instead of the bell-shaped GRP distribution shown in Figure 4(d), the memory processes that underlie this commercial and brand name rather suggest a uniformly spaced, dripping schedule. Significant increases were also found for the brands shown in Figures 5(a) and 5(c).

There is obviously a strong relation between the parameter values of a fitted-recall probability function and the type of optimal advertisement scheduling, which can be classified as follows. For brands that only consist of a stationary base rate process (i.e., no additional encoding during the campaign), dripping schedules with small amounts of GRP will be most cost effective. This situation arises three times for impact and 15 times for brand name recall. Brands that are memorized according to a one-store model with strong forgetting or that suffer from strongly saturated learning benefit most from a spaced distribution of GRP (dripping). Brands that are learned without saturation and that show very slow forgetting afterwards benefit most from a single burst at the beginning of a campaign (bursting). The same holds when the recall threshold increases (new brands or campaigns). In this way, the increased impact will

**Table 3 Implications of Different Model Versions for Scheduling**

| Base rate | Learning | Forgetting | Recall threshold | Schedule |
|---|---|---|---|---|
| Exists | Very slow | | | Drip |
| | | Rapid | Low | Drip |
| | Saturated | | Low | Drip |
| | Not saturated | Very slow | | Burst |
| | Not saturated | Moderate, no consolidation | Low | Burst, then drip |
| | Not saturated | Rapid, with consolidation | Low | Burst, then drip |
| | | | High | Burst |

fall within the campaign period as much as possible. Other situations give rise to a mixture of bursting and dripping. Examples are brands and commercials that are memorized according to a two-store model without saturated learning, with strong forgetting in the first store, and with slow or no forgetting in the second store. Here, on the one hand, the memory consolidation process of the second store favors a single burst at the start of the campaign, but the strong forgetting in the first store requires a uniform distribution of GRPs over the campaign weeks. A mixture of bursting and dripping also applies to one-store models with moderate forgetting. Examples of the abovementioned brands and optimal GRP distributions are shown in Figure 5. Table 3 summarizes the above implications of the model for optimal advertisement scheduling.

## Discussion

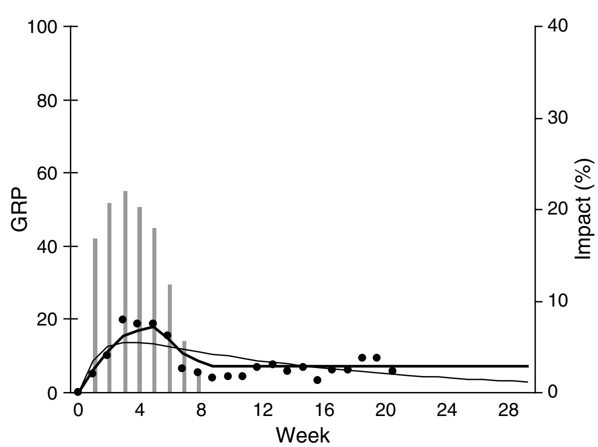
In this paper, we presented and applied a neurobiologically informed model to the recall of advertising material as a function of multiple exposures to the same advertisement (print or TV). To our knowledge, it is the first application of neurobiological process models to advertising. The same processes are involved in a wide range of learning and forgetting situations to which our model has been fitted successfully. This study shows that the model also captures the learning and forgetting of advertising material.

An important and essential feature of the model is its ability to describe situations with multiple learning trials, which was demonstrated in previous studies with model fits to psychological laboratory data, and in the present setting to recall data for multiple advertisement exposures. About 93% of the fitted recall-probability functions were not rejected. This shows that the recall functions, which are derived at neural and individual levels, give good approximations of the impact and brand name recall data at the aggregate level. Two reasons could explain this: (1) the variation in the number of advertisement exposures among respondents in an advertising week is rather small; (2) the memory processes may be quite homogeneous over respondents. Departures from these properties will imply an overestimation of the recall functions used here at aggregate level, because recall probability is a concave function of memory intensity. Explicit modeling of variations over respondents should reveal the precise amount of overestimation. Because the overestimation applies to the entire course of the recall-probability function, we expect that the effects on the optimal schedule types shown in Table 3 will be limited.

The fits of our model to the SPOT data have shown that the feed-forward structure of the notion of memory stores, by which we allow memories to consolidate to one or more stores, is an essential element in the applicability of our model. There are numerous instances in the SPOT data that show a decline of impact and brand familiarity during advertising, which then stabilize at constant values after a campaign. From our model perspective, these cases all result from a rapid decline in the first memory store, from which a part of information is consolidated to a second store.

As a comparison, we implemented and applied the model proposed by Naik et al. (1998) to the SPOT data. Their model gave considerably poorer fits with as many free parameters as in the memory chain model, with about 22% of the fits rejected (Table 2). The poorer fits were caused to a large extent by the model rejections for situations described in the previous paragraph. The use of one forgetting parameter ($\delta$ in Naik et al. 1998) is not sufficient to capture complex behaviors of advertising effectiveness during and after campaigns. Naik et al.'s model must fit variations in memory decline during and after a campaign with the same value for the forgetting parameter $\delta$. This leads to overly smooth fits, and an inability to

**Figure 6 Fits of a Two-Store Memory Chain Model (Thick Line) and of the Model of Naik et al. (Thin Line), with the Same Number of Parameters, to Impact Data (Dots) of a Meal Sauce Brand**




*Notes.* The model of Naik et al. is rejected at $\alpha = 0.05$, while the 2-store model is not rejected at very large test sizes ($\alpha > 0.85$).

handle changes in memory decline over time due to, for instance, long-term memory consolidation (see Figure 6). This is also true for other models, such as the Vidale-Wolfe and Nerlove-Arrow models and Brandaid (see the paper by Naik et al., where expressions for different models are listed in Table 2).

Our cross-validation study showed that the parameter estimates are stable over time. The models thus used to optimize advertisement scheduling showed that optimal schedules fall in one of three types: (1) a GRP burst at the start of the campaign, (2) a uniform, dripping GRP distribution, (3) a combination of bursting and dripping, with the greatest GRP expenditure at the start of the campaign and a more even distribution of GRP in later weeks. The implications of the memory model parameters for optimal scheduling, as shown in Table 3, could thus be used to target advertising to different groups of products or brands (see also Iyer et al. 2005).

The optimal schedules that we obtained are all decreasing or constant in GRP as a function of time. This does not mean, however, that the schedules implied by our model are all monotonic (e.g., decreasing). Further model analyses have shown that optimal GRP schedules may also be nonmonotonic. The critical parameters in this situation are the decline parameters of the memory stores and the rate of induction between the first and second stores. Nonmonotonic optimal schedules arise when the induction rate is larger than the memory decline rate of the first store. This situation never arose in our fits.

The contribution of this study from a scheduler's or manager's point of view is the effect of advertisement memory parameters on optimal scheduling. Decisions concerning the planning of advertisement campaigns could be supported better when knowledge is available about relations between memory parameters and advertisement characteristics. Although finding these relations is left as a challenge for future research, we could make several hypotheses here. For instance, we believe that a small saturation of learning parameter $r_{\max}$ is related to a more or less worn-out advertising concept, that encoding $\mu_1$ and learning rate $v$ are highly influenced by personal involvement and attention, and that the recall threshold $b$ is influenced by novelty and the coherence of an advertising message (i.e., small for well-known or coherent messages). Answers to these open questions (e.g., through psychological experiments) could lead to a better and more complete understanding of the relation

$$\textit{ad properties} \rightarrow \textit{memory parameters} \rightarrow \textit{scheduling}.$$

In future research, we will extend our model with a decision-analytic component for consumer brand choice and buying behavior (e.g., see Akçura et al. 2004). We will also consider explicit modeling of competition effects. In this respect, it could be interesting to integrate the model part on copy wearout in Naik et al. (1998) with our model.

**Acknowledgments**
This research was supported by NWO, the Netherlands Society for Scientific Research. The authors are very grateful to Jan Ligthart for providing the SPOT data and to Giep Franzen for many helpful discussions. They also thank the reviewers for their many useful and detailed comments.

## References

Abeles, M. 1991. *Corticonics*. Cambridge University Press, Cambridge, UK.

Akçura, M. T., F. F. Gönül, E. Petrova. 2004. Consumer learning and brand valuation: An application on over-the-counter drugs. *Marketing Sci.* **23** 156–169.

Atkinson, R. C., R. M. Shiffrin. 1968. Human memory: A proposed system and its control processes. K. W. Spence, J. T. Spence, eds. *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 2. Academic Press, New York.

Baddeley, A. D., D. J. A. Longman. 1978. The influence of length and frequency of training session on the rate of learning to type. *Ergonomics* **21** 627–635.

Bahrick, H. P., L. E. Bahrick, A. S. Bahrick, P. E. Bahrick. 1993. Maintenance of foreign language vocabulary and the spacing effect. *Psych. Sci.* **4** 316–321.

Barnes, C. A., B. L. McNaughton. 1980. Spatial memory and hippocampal synaptic plasticity. D. Stein, ed. *The Psychobiology of Aging: Problems and Perspectives*. Elsevier/North Holland, New York, 253–272.

Burtt, H. E., E. M. Dobell. 1925. The curve of forgetting for advertising material. *J. Appl. Psych.* **9** 5–21.

Chessa, A. G., J. M. J. Murre. 2004. A memory model for Internet hits after media exposure. *Physica* A **333** 541–552.

Daley, D. J., D. Vere-Jones. 1988. *An Introduction to the Theory of Point Processes*. Springer-Verlag, Inc., New York.

Diggle, P. J. 1983. *Statistical Analysis of Spatial Point Patterns*. Academic Press, London, UK.

Glenberg, A. M. 1976. Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *J. Verbal Learn. Verbal Behav.* **15** 1–16.

Glenberg, A. M., T. S. Lehmann. 1980. Spacing repetitions over 1 week. *Memory Cognition* **8** 528–538.

Goldman-Rakic, P. S. 1992. Working memory and the mind. *Sci. Amer.* **267** 111–117.

Goldman-Rakic, P. S. 1995. Cellular basis of working memory. *Neuron* **14** 477–485.

Iyer, G., D. Soberman, J. M. Villas-Boas. 2005. The targeting of advertising. *Marketing Sci.* **24** 461–476.

Janssen, S. M. J., A. G. Chessa, J. M. J. Murre. 2005. The reminiscence bump in autobiographical memory: Effects of age, gender, education, and culture. *Memory* **13** 658–668.

McGaugh, J. L. 2000. Memory—A century of consolidation. *Science* **287** 248–251.

Meeter, M., J. M. J. Murre. 2004. Consolidation of long-term memory: Evidence and alternatives. *Psych. Bull.* **130** 843–857.

Meeter, M., J. M. J. Murre. 2005. TraceLink: A model of consolidation and amnesia. *Cognitive Neuropsych.* **22** 559–587.

Merzereau, P., L. Battais. 2000. Continuity or burst? How do TV strategies compare? *Admap* (October) 30–35.

Murdock, B. B., Jr. 1961. The retention of individual items. *J. Experiment. Psych.* **62** 618–625.

Murdock, B. B., Jr. 1974. *Human Memory: Theory and Data*. Lawrence Erlbaum Associates, Inc., Potomac, MD.

Murre, J. M. J. 1996. TraceLink: A model of amnesia and consolidation of memory. *Hippocampus* **6** 675–684.

Murre, J. M. J., K. S. Graham, J. R. Hodges. 2001. Semantic dementia: Relevance to connectionist models of long-term memory. *Brain* **124** 647–675.

Murre, J. M. J., M. Meeter, A. G. Chessa. Modeling amnesia: Connectionist and mathematical approaches. M. J. Wenger, C. Schuster, eds. *Statistical and Process Models for Neuroscience and Aging*. Erlbaum, Mahwah, NJ. Forthcoming.

Naik, P. A., M. K. Mantrala, A. G. Sawyer. 1998. Planning media schedules in the presence of dynamic advertising quality. *Marketing Sci.* **17** 214–235.

Peterson, L. R., M. J. Peterson. 1959. Short-term retention of individual verbal items. *J. Experiment. Psych.* **58** 193–198.

Rubin, D. C., A. E. Wenzel. 1996. One hundred years of forgetting: A quantitative description of retention. *Psych. Rev.* **103** 734–760.

Rumelhart, D. E. 1967. The effects of interpresentation intervals on performance in a continuous paired-associate task. Technical Report 116, Psychology Series, Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford, CA.

Simon, J. L. 1979. What do Zielske's real data really show about pulsing? *J. Marketing Res.* **16** 415–420.

SPOT. 1998. The effectiveness of TV-advertising. Internal report. (Available at www.spot.nl/onderzoek/onderzoek/downloads/Effectiviteit.pdf, in Dutch: *De effectiviteit van tv-reclame*), The Foundation for Promotion and Optimization of Television Advertising (SPOT), Amsterdam, The Netherlands.

Stoyan, D., W. S. Kendall, J. Mecke. 1987. *Stochastic Geometry and Its Applications*. John Wiley Sons/Akademie Verlag, Berlin, Germany.

Talamini, L. M., M. Meeter, J. M. J. Murre, B. Elvevåg, T. E. Goldberg. 2005. Integration of parallel input streams in parahippocampal model circuits. *Arch. General Psychiatry* **62** 485–493.

Wickens, T. D. 1982. *Models for Behavior: Stochastic Processes in Psychology*. Freeman, San Francisco, CA.

Zielske, H. A. 1959. The remembering and forgetting of advertising. *J. Marketing* **23** 239–243.