Foundations and Trends® in
Marketing
Vol. 1, No. 4 (2006) 231–320

DOI: 10.1561/1700000011

Eye Tracking for Visual Marketing

Michel Wedel[1] and Rik Pieters[2]

[1] *Robert H. Smith School of Business, University of Maryland, 3303 van Munching Hall, College Park, MD 20742-1815, USA, mwedel@umd.edu*

[2] *Marketing Department, Tilburg University, 5000 LE Tilburg, The Netherlands, pieters@uvt.nl*

Abstract

We provide the theory of visual attention and eye-movements that serves as a basis for evaluating eye-tracking research and for discussing salient and emerging issues in visual marketing. Motivated from its rising importance in marketing practice and its potential for theoretical contribution, we first review eye-tracking research for visual marketing. Then, we discuss the structure of the eye, the visual brain, eye-movements, and methods for recording and analyzing them. Next, we describe our theory and review eye-tracking applications in marketing based on it. We conclude with an outlook on future theory and method development and recommendations for practice.

1

Introduction

The importance of auditory, gustatory, olfactory, and kinaesthetic stimuli in marketing to consumers can hardly be overstated, but visual stimuli have dominated research, and people are primarily visually oriented. As consumers, we are exposed every day to several hundreds of advertisements on television, in newspapers, magazines, the yellow pages, retail feature ads, and on internet sites, and we experience even more, implicit ads in the form of product packages in stores and at home, visual messages on the side of trucks, road signs, food wrappers in restaurants, on service provider uniforms, t-shirts, CDs, and electronic devices. And the eyes guide consumers across shelves, through aisles, stores, malls, and websites when exploring, searching, and making decisions on products and brands. In all these situations "visual marketing" is relevant, that is, the strategic utilization by firms of commercial and non-commercial visual signs and symbols to communicate with consumers in order to establish and maintain mutually profitable relationships. If indeed "seeing is believing," and "believing is buying," it is important to manage closely what consumers see, and to track this to optimize profitability. This is increasingly recognized in business practice.

In the last decade there has been a rapid growth in commercial applications of eye-tracking technology in the United States, Europe, Asia, and Australia to assess the effectiveness of visual marketing efforts. Firms such as Kraft Foods, Microsoft, Google, Yahoo, IBM, Pepsico, Pfizer, P&G, and Unilever are leading users of the methodology in product and communication development, and in pre- and post-tests of their visual marketing activity. With the increasing demand from marketing practitioners, global providers of eye-tracking data have emerged, conducting hundreds of studies each year. Such commercial research companies include, for example, Colmar Brunton, Eye-Square, the Japanese Institute of Consumer Marketing Research, LC Technologies, the Pretesting Company, Perception Research Services, the Siegfried Vogele Institut for Direct Marketing, Verify International, VisionTrack SR-Research, and Tobii.[1] The growth of eye tracking is partly driven by technological innovations in the development of unobtrusive and precise eye tracking devices and sharp declines in the costs of these devices. Until recently, the commercial use of eye tracking was limited, because it was cumbersome for the participants, time consuming for researchers, and expensive. Moreover, erroneous beliefs in marketing academia and practice about the role and importance of stimulus- and memory-based processes and their interactions hindered progress.

That is, in marketing the study of visual attention through eye tracking was hampered by the view that attention is only a gate through which information enters on its ways to higher-order cognitive processes of more interest, that gaining and retaining attention is easy through contrast and repetition, and that measuring attention with eye tracking is difficult. Moreover, theories of consumer decision making were strongly memory-based, emphasizing the off-line processes well after information intake, rather than the on-line, moment-to-moment

[1] Colmar Brunton (www.cbr.com.au/), Eyetracking.com (www.eyetracking.com), Eye-Square (www.eyesquare.com/), the Japanese Institute of Consumer Marketing Research (www.jmrlsi.co.jp/english/), LC Technologies (www.eyegaze.com), the Pretesting Company (www.pretesting.com/), PRS (www.prsresearch.com/), Psyma (www.psyma.com). The Siegfried Vogele Institut for Direct Marketing (www.sv-institut.de), Verify International (www.verify.nl), VisionTrack (www.visiontrack.com), SR_Research (www.sr-research.com) and Tobii (www.Tobii.com).

processes that co-occur with and are reflected in eye movements across the marketing stimuli of interest. More generally, visual and perceptual processes were neglected in marketing research in favor of a focus on higher-order cognitive processes. That is, marketing theory and practice assumed that information acquisition and decision making were temporally separated, that information acquisition was not a key challenge, and that decision making was rooted in more or less complex manipulations of abstract constructs, that could be assessed by consumers' verbal self-reports and choices. It has become increasingly clear that such views are flawed.

That is, eye movements are tightly coupled with visual attention which makes them eminent indicators of the covert visual attention process. Psychological research reveals that visual attention is not only a gate, as suggested by hierarchical processing models such as AIDA (Strong, 1920; Starch, 1923), but reflects higher-order cognitive processes (Rizzolatti et al., 1994) and is closer to actual behavior than intuition informs us (Russo, 1978; Steinman, 2004). Gaining and retaining attention is difficult, because it is difficult to break through high levels of visual clutter in various media (Burke and Srull, 1988; Keller, 1991; Kent, 1993; Mulvihill, 2002; Schwartz, 2004). Finally, measuring visual attention is now easy with modern eye-tracking equipment. Bettman et al. (1991, pp. 74–75) already lauded the advantages of eye tracking over other methodologies to gain detailed insights into moment-to-moment ad processing and consumer decision making, but pointed out the — then — operational difficulties that prevented the technology to reach its full potential for marketing. This situation has changed in recent years due to new generations of infra-red eye trackers, which enable eye-movement recordings for large quantities of stimuli and consumers under natural exposure conditions at high precision and low cost.

With the current ease of eye-movement recordings and the emergence of a body of theory of visual attention and stimulus-based decision making,[2] the door is open for further research on visual

[2] Here and in the sequel the term stimulus-based decision making obviously does not preclude memory-based processes. Stimulus- and memory-based decision processes jointly

marketing, building on and extending what has already become known in recent years through eye-tracking research. These developments not only provide richer opportunities for marketing practitioners, but also for rigorous academic research on the value of such practices. The history of academic research in this area started in the early 1900s, when Nixon (1924), Poffenberger (1925) and Karslake (1940) applied eye-movement research to determine the attention capture value of magazine and newspaper advertisements with varying sizes, and color and black-and-white ads. After a period of relative silence, in the 1970's there was a revival of the interest in the methodology through the work of Van Raaij (1977), Russo (1978), Treistman and Gregg (1979), and Kroeber-Riel (1979). The 1990's and 2000's have seen a surge of the interest in eye tracking, in part driven by advances in technology. Now a sizable and growing body of literature exists on attention to visual marketing stimuli, including out-door advertising, point-of-purchase material, print ads, catalogues, DM letters, television commercials, shelves, web-pages, and yellow page ads.

Recently, summaries of eye tracking research were provided by Duchowski (2002), and Wedel and Pieters (2007). This survey goes beyond these summaries by (1) providing the foundations of visual attention and eye tracking, (2) providing a conceptual framework for eye-tracking research in marketing, and (3) reviewing the marketing literature within that conceptual framework. We begin with a review of the anatomy of the eye and the visual brain, and of visual and attentional processes.

shape consumer choice. Memory-based decision processes, however, do not require direct stimulus-based input and processes, and we strive to open new roads to understanding the latter in the present work.

2

The Eye and Its Movements

Light enters the human eye through the cornea (the outermost layer of the eye) and the lens (and two chambers with fluids), and falls on the retina at the back of the eye. The lens refracts light and focuses it on the retina, reversing the image. The curvature of the flexible lens is changed by ciliary muscles. When these contract, the lens attains a round shape, which enables focus on close objects. When the ciliary muscles are relaxed, the shape of the lens flattens, which enables focus on distant objects. The iris, between the cornea and the lens, consists of a ring of muscles and the pupil is in its center. The size of the pupil — ranging from about 2 mm to about 8 mm under normal conditions — is adjusted by these muscles to maintain a constant level of light on the retina under various conditions. It dilates in dimmer and constricts in brighter light, dilates for a wider and constricts for a narrower visual angle, and dilates for a smaller and constricts for a larger depth of focus. Changes in pupil diameter (intra-ocular movement) is also thought to reflect information load and affective processes, as discussed later.

The retina contains two types of photoreceptors, rods and cones. Rods (about 120 million) are more light sensitive than cones, but insensitive to color. The retina contains other types of cells as well that mainly serve to transport signals: horizontal cells that collect signals from several cones, amacrine cells that do the same for (peripheral) rods, bipolar cells that connect the rods and cones to the ganglion cells that are used to send the output from the eye to the brain through the optic nerve. The place where the optic nerve exits the eye is called the blind spot, because there is no vision at that location. Cones (6 to 7 million), providing the eye's color sensitivity, are mostly concentrated in the central yellow spot in the retina known as the macula. In the center of the macula, the fovea centralis (fovea, for short) is a 0.3 mm diameter rod-free area with very thin, densely packed cones. The fovea is directly opposite to the lens and enables high visual resolution[1] in the center of the visual field (which is about two degrees of visual angle, the size of a thumbnail at arm's length). Acuity drops off rapidly to the periphery, due to an increase in the size and corresponding decrease in the density of the photosensitive cells in the retina. This composition of cells, however, makes those regions of the retina sensitive to large objects and in particular to motion. There are an estimated 120 milion rods and cones, but only about 1.5 milion ganglion cells, which makes a very large reduction in the information collected by the retina necessary. This reduction is achieved by low visual acuity in the periphery. The retina is highly specialized: the fovea is sensitive to high spatial frequencies (local, fine visual detail such as the edges of objects), but has slower signal transmission, the periphery is sensitive only to low spatial frequencies (global, coarse visual information such as large patches of luminance contrast), but has a high speed signal transmission. Ganglion cells in the fovea integrate information from cones covering about 0.03 degrees of visual angle resulting in high acuity, while these cells integrate information from rods in the periphery covering about 3 degrees of visual angle, which results in low acuity. Figure 2.1 shows the human eye with several of these components labeled. For a detailed review of the human eye, see Palmer (1999) or the Physiology of the Senses Website by Tutis Vilis.[2]

[1] Resolution is measured in cycles per degree of visual angle, the resolution of the fovea is 12 cycles per degree.
[2] http://www.physpharm.fmd.uwo.ca/undergrad/sensesweb/, last accessed on March 14, 2008.

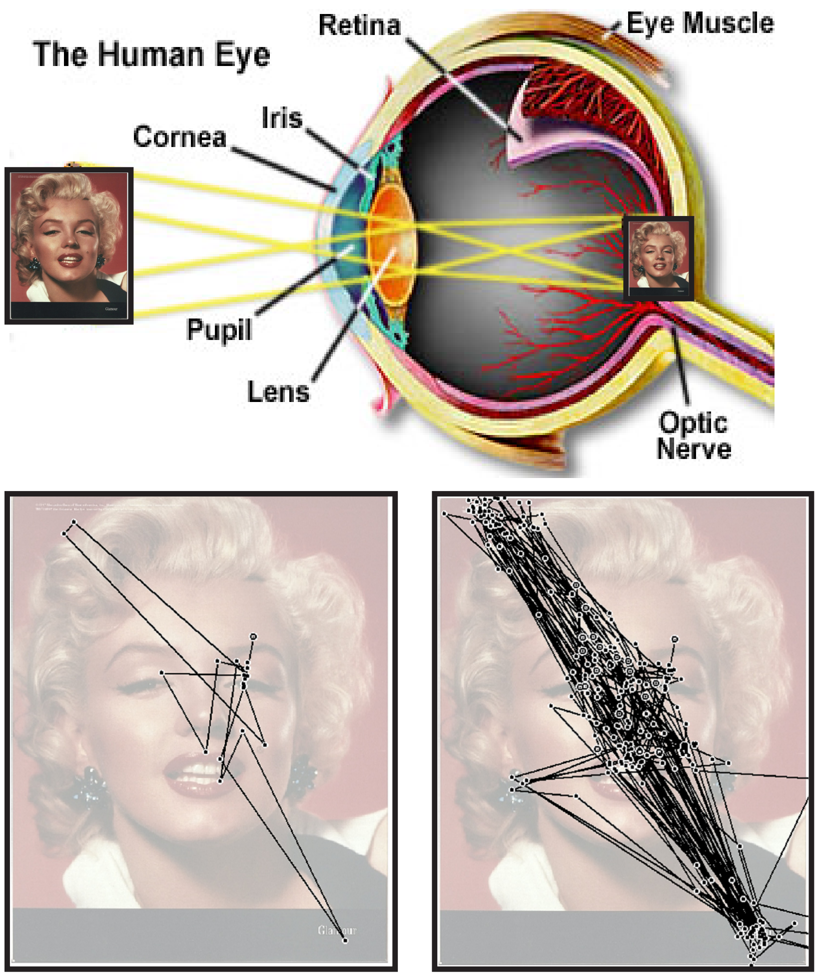

Fig. 2.1 The human eye and eye movements on a Mercedes ad (Top-left = original ad, bottom-left = scan path of single consumer, bottom-right = scan paths of 20 consumers, top-right = brand, logo, and key word; Dots are fixations, lines are saccades). Mercedes logo is hidden on the right cheek of Marilyn.

At any given point in time only about 8% of the visual field is projected on the fovea, and thereby available for detailed processing. The remainder of the visual field, i.e., the parafovea and periphery, is apprehended in a coarser, mostly non-colorful and blurred fashion. Parafoveal and peripheral vision for fine detail and color are very limited. The limited range of high acuity vision across the retina makes it necessary to move the eyes in order to successively project specific small parts of scenes on the fovea. These (extraocular) eye movements are controlled by six muscles attached to the eye, and that cause upward/downward (superior/inferior rectus), inward/outward (medial/lateral rectus), and downward/upward rotational movements away from the midline (superior/inferior oblique). Any given eye-movement may involve several of these muscles.

We generally have the experience that the complete visual field is clear and available at a high resolution. This illusion of "complete vision" is due to the fact that whatever we focus on in a scene is clear. Whenever we move our eyes to focus on something else vision is suppressed, but the new location or object is clear again. This gives rise to the impression that much detail of the scene can be perceived parafoveally and peripherally. In addition, we are inclined to believe that we can readily and easily move attention to some part of the visual field, without moving the eyes, that is, by looking "from the corner of our eyes." Both of these ideas are incorrect, but persistent. We are aware of the fact that we look at informative locations in scenes, but introspection falls short of informing us precisely how, when and to what we move our eyes.

Dodge (1900) was one of the first to reveal that what we believe to be smooth movements of our eyes in fact consists of sequences of fixations and saccades — the illusion of "smooth eye-movement." Saccades are ballistic jumps of the eyes serving to project specific locations of the scene onto the fovea. The saccade is the fastest movement in the human body, with angular speeds of up to 1000 degrees/second; humans make around one hundred and seventy thousand saccades a day. They are often prompted by peripheral vision and typically last 20–100 milliseconds. Vision, specifically of high spatial resolutions, is suppressed during a saccade. During a fixation, which typically lasts around 200–500 millisecond (Rayner, 1998), a contiguous area of the scene is projected onto the fovea for detailed visual processing. Yet, the eye is never completely still and moves even during a fixation. These movements include microsaccades, tremor and drift, the first being very small, about 0.2 degrees at most, and imperceptible movements that serve to continuously stimulate the cells in the retina, without which the visual percept would quickly fade. Vergence eye movements are movements where both eyes turn inward, in order to keep an object that moves toward or from the observer in focus. Finally, smooth pursuits are eye movements to follow moving objects, with speeds of up to 100 degrees per second. Smooth pursuits cannot be made voluntarily, without the presence of a moving object.

Figure 2.1 shows the attention scan path of a single (bottom-left) and 20 consumers (bottom-right) across a magazine ad for Mercedes (top-left). In the ad, the brand name is depicted very small at the top, and the brand logo is hidden in Marilyn Monroe's beauty mark on her left cheek. Of note are the non-random, systematic scan paths across consumers, demonstrating how eye movements patterns provide insight into "information intake" resulting from perceptual and cognitive processes to which we return later.

3

The Visual Brain

The signals from the retina are transported by the ganglion cells through the optic nerve to the visual cortex in the back of the brain, via the Lateral Geniculate Nucleus (LGN, part of the thalamus, another part of the information goes directly to the Superior Colliculus, discussed below). The LGN is a layered structure. Some layers have large (magnocellular) cells that take a relatively short time to process mostly global, coarse information from the periphery of the visual field, and are part of a visual pathway that deals with location and motion. Other layers have smaller (parvocellular) cells that process mostly local, detailed information from the fovea about object specifics and text, more slowly.

The visual cortex consists of a number of areas that are labeled V1 to V5, occupying about one-third of the cerebral cortex. These layers in the visual cortex are hierarchically organized, with lower levels tuned to more basic features, and higher levels tuned to more complex patterns, structures, and objects. Many of these areas, especially the lower visual areas, contain retinotopic maps of the visual field, where any given location in the map corresponds uniquely to a location in the visual field. The area V1 receives information directly from the LGN and transmits information to two neural pathways (Ungerleider and Mishkin, 1982). The ventral ("what") pathway begins in V1, and goes through V2, ventral V3 and V4 into the Inferotemporal Cortex (ITC). This pathway is involved in form and object recognition (faces, people, places) and long term memory. The dorsal ("where") pathway begins in V1, and goes through dorsal V3 and V5 into the Posterior Parietal Cortex (PPC). This pathway is involved in the representation of object locations and motion, and the guidance of hand and eye movements (Ungerleider and Mishkin, 1982). The two pathways share information, which enables the localization of objects and their movements in the visual field.

V1 neurons act as a selective spatiotemporal filter, with cells arranged in six layers responding to local contrasts in specific basic perceptual features including luminance, orientation, motion, direction, spatial frequency, and color. The mapping between a given location in the visual field and a location in V1 is very precise, but the foveal region occupies more space than the peripheral regions. V2 receives input from V1, has feedback connections to that area as well, and shares many of its properties. It is tuned to features such as the orientation, spatial frequency, and color of objects, but more so than V1 to moderately complex patterns. It is involved in figure-ground segmentation and perceptual grouping, which are important in object recognition. V3 receives input from V1 and V2 and projects to the PPC. It may contain a representation of the entire visual field and play a role in global motion perception. V3 is part of both the dorsal and the ventral stream. V4 receives input from V2 and sends connections to the ITC. It is the first area that shows strong influences of attentional focus. It is tuned to object features of intermediate complexity, and is involved in form recognition and color vision. V5 (MT) processes both motion of complex visual features and shapes and integrates local visual motion into object-motion. Attention focuses the response of cells in MT. V3 projects to the PPC through MT. MT connects to the area MST where different areas specialize in sensing motion of objects in the visual field and flow patterns of the entire field caused by movement of the individual. MT connects to the superior colliculus (SC) and the frontal eye-fields (FEFs).

The SC and FEF are involved in guiding eye movements, and both hold a topographical representation of the visual field, with an essentially one-to-one spatial representation of locations in the visual field to (clusters of) cells in these brain structures. About 20% of the nerve pulses from the retina bypass the LGN and are relayed directly to the SC, as part of the retinotectal pathway, that projects onto the PPC, amongst others. The SC responds to salient stimuli in the periphery of the visual field and directs short latency saccades to them.[1] The frontal eye-fields (FEFs) direct longer latency saccades to targets that are held in working memory in the prefrontal cortex (PFC). Areas of the SC and FEF involved in long amplitude saccades are part of the dorsal visual stream that passes from the peripheral retina into the parietal cortex, while those producing short amplitude saccades are part of the ventral stream, that passes from the fovea through the inferotemporal cortex (Schiller, 1998; Thompson, 2005).[2] The parietal cortex contains a retinotopic representation of salience of spatial locations, used for eye- and hand-movements (the Lateral Intraparietal area LIP), integrates visual, auditory, and somatosensory information (the Ventral Intraparietal Area VIP) and maps perceived objects into body coordinate positions (the Medial Intraparietal area MIP).

The design of the eye and visual brain are tailored to specific tasks. The fovea is involved in (slow) intake of local, detailed, colorful information from a small area in the scene, the information being relayed via the LGN to other parts of the visual brain for further processing of forms, shapes, and objects, while the periphery of the retina, the SC is involved in the very fast and large eye movements to change the point of regard and bring new locations into focus. The areas in the visual brain are highly specialized to process information collected during eye-fixations, and continuously interact with areas that direct eye movements to salient and/or informative locations in visual scenes and stimuli, as well as to higher cortical areas such as the frontal cortex, which enables

[1] Saccadic latency is the time before a next saccade is initiated, which is by and large the same as the fixation duration.

[2] Saccadic amplitude is the size of a saccade, expressed in degrees of visual angle or spatial distance across the image.

purposeful and goal-directed eye movements. The recording of these eye-movements therefore enables inferences on information acquisition and processing from visual marketing stimuli. For a detailed review of the visual brain, see Fuster (2003), wikipedia[3] or the Physiology of the Senses Website by Tutis Vilis.[4]

[3] http://en.wikipedia.org/wiki/Visual_cortex, last accessed on March 14, 2008.
[4] http://www.physpharm.fmd.uwo.ca/undergrad/sensesweb/, last accessed on March 14, 2008.

4

Eye-Movement Recording

There are three common ways to record eye movements with high-levels of accuracy (Young and Sheena, 1975; Duchowski, 2003): (1) EOG, electro-oculography based, which measures electric potential differences around the eyes, which vary with rotation, (2) SCL, scleral contact lens based methods, which measure eye movements relative to head position, using a wire coil in a contact lens inserted in the eye, and (3) VIROG, videobased infrared oculography. The latter method, termed “infrared eyetracking,” has become more popular for eye tracking in market research and other applied settings, and will be described in some more detail.[1] A more extensive description of other eye tracking methodologies is provided in Duchowski (2003).

Infrared light is emitted from a light source and reflected of various layers in the eye, thereby creating a series of four, so called, Purkinje reflections on the eye, respectively from the front and back of the cornea and lens. Although several of the reflections can be measured

[1] In early accounts, eye movements were sometimes observed directly through a one-way mirror or specially prepared viewing wall (e.g., Berlyne, 1958). There are promising applications (e.g., Shimojo et al., 2003) with regular video camera recording for eye tracking. We are not aware of comparisons of the spatial resolution of these measurements with those of other methodologies.

simultaneously to more precisely measure the "point of regard" (POR), i.e., the location in the scene that is projected on the fovea, head stabilization of the participants is required by means of a mouthbit, and/or forehead or chinrest. Therefore, typical commercial eye-tracking devices measure only the first Purkinje reflection off the front of the cornea of a single or of two eyes, which provides sufficient accuracy. The difference between the pupil center and the corneal reflection changes with eye rotation, but remains relatively constant with normal head movements. This feature is useful because it allows fairly free head movements during recording, and thus more unobtrusive measurement. The POR is determined from the angle and distance of the corneal reflection (the "glint") from the pupil center, after a calibration task. Such infrared eye tracking typically has a temporal resolution of 50 hertz (Europe: 20 milliseconds) or 60 hertz (US) (or multiples of this), and spatial resolution of 0.5°, which is sufficiently accurate for commercial and academic applications in marketing.

From the raw gaze patterns, eye fixations on the visual stimulus and the saccades between them are identified as key ("extra-ocular") eye-movement measures. The diameter of the pupil is often obtained as a corollary of infrared eye tracking as an additional ("intra-ocular") eye-movement measure. Often, anomalies caused by blinks and tear fluid will have to be filtered out of the raw gaze pattern.

Salvucci and Goldberg (2000) classify the algorithms that are available to determine eye fixations in two general classes: spatial and temporal methods. Spatial methods define fixations in terms of velocity, dispersion, or area-of-interest. Temporal methods do that in terms of duration and local adaptation. Salvucci and Goldberg compare several algorithms and conclude that velocity-based and dispersion-based algorithms perform similarly, and perform quite well. Area-based algorithms are often restrictive and can generate biases. The use of temporal information helps the identification of fixation sequences, and adaptive algorithms generate robust results, even in the presence of noise in the fixation data. In practice methods are often used that specify thresholds for distance (small distances identifying a fixation), velocity (high velocities defining a saccade), or a combination of the two are robust, and in general perform well.

Various companies in the United States and Europe, such as ASL, SMI (iView), and SRI Research (EyeLink2),[2] offer table-mounted systems and light-weight head mounted eye-tracking systems, the latter being particularly useful when eye recording during unrestricted movement of people, such as during (simulated) shopping tasks, is paramount. Tobii Systems[3] offers an eye-tracking system that integrates eye-tracking cameras in the rim of a 1280×1024 pixel LCD monitor (17 inch or larger). Figure 4.1 shows a graph of the raw x- and y-coordinates of the gaze samples, with fixation locations as identified with a standard distance/duration threshold algorithm (20 pixels, 30 ms) superimposed. Revised ideas about the role of stimulus-based processes in decisionmaking, as well as the comparatively low costs of new generations of eye tracking systems, short calibration times, natural exposure conditions, unobtrusive measurement, the possibility

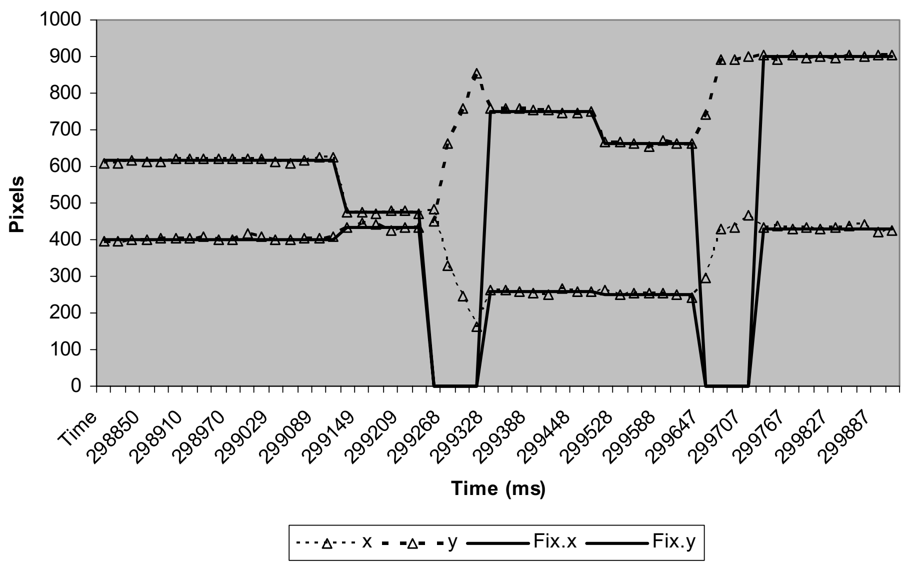

Fig. 4.1 Sample eye tracker data collected with Tobii 1750. The x- and y- coordinates of the raw gaze samples are shown, with fixation positions identified with standard algorithms superimposed.

[2] www.a-s-l.com; www.smivision.com; www.sr-research.com
[3] www.Tobii.com.

to track consumers with eyeglasses and contact lenses, and improved processing software have all contributed to the exponential growth of eye-movement applications in marketing practice, and their use for theory development and testing in academic research. This is illustrated by over 600,000 hits that Google returns early 2008 under the search terms "eyetracking and marketing."

5

Eye-Movement Metrics and Inference

Datasets produced in eye-tracking studies can become very large. Consider, for example, a study in which 100 participants are exposed to 25 print ads with an average exposure of 4 seconds per ad. With a measurement frequency of 50 hertz per second (or 60 in the United States), the eye-tracking equipment makes video snapshots, called samples, of the x and y location of the POR. This would lead to, respectively 1,000,000 and 1,200,000 (US) datapoints for the (extra-ocular) eye movements alone. Dealing with that amount of data is nontrivial, and theory on the underlying perceptual and cognitive processes is needed to formulate predictions and make sense of the data. In some instances, the individual samples of the POR measurements are used for inspection, but current software automatically aggregates the raw data over time into sequences of fixations and saccades. Here, we provide some basic principles for deriving metrics from such eye-movement data.

Each fixation has x–y coordinates (for instance in screenpixels), a rankorder (from first to last during an eye-movement recording) and a duration (in milliseconds). An eye-movement recording of a particular individual thus comprises a sequence of fixations, each of a particular duration at a specific location in the scene, with saccades between them.

Aggregation can take place across the temporal and spatial components of sequences of eye movements, and across participants, leading to a variety of metrics. The number of possible metrics that can be derived from eye movements is large, and a host of measures have been proposed and are being used in psychology, engineering, and marketing (see Inhoff and Radach, 1998; Jacob and Karn, 2003; Rayner, 1998). Here, we distinguish — in a first step toward a more formal framework — sequence, transition, distance, density, and duration measures, across x–y coordinates, grid cells, meaningful objects in the scene, the entire scene itself, and participants. The resulting metrics each aggregate across one or more of the dimensions of the eye-tracking data — time, space, and individuals — to focus on specific aspects. *Time* covers the duration and sequence of eye-movement events. *Space* covers the locations and objects in specific stimuli or across multiple stimuli where eye-movement events take place. *Individuals* in the sample of participants are allocated to specific stimulus or task conditions to introduce systematic heterogeneity in eye movements or whose natural heterogeneity is accounted for.

5.1 Time

If interest is in the processing order of the scene, *sequence and transition measures* are used such as (1) the location or area of the first fixation, (2) the order in which locations in the scene are fixated, (3) the time until an object, element or location is fixated first, (4) the time from the first until the final fixation on such a location during the first visit (first pass time), and (5) the switching frequencies or probabilities between specific locations in the scene. The complete nth order fixation and saccade sequence can be examined without aggregation, and aggregation is at a maximum when only 1st order transitions between locations are examined. The time until the first fixation has been used to study the visual pop-out or "lift" of elements and objects (Chandon et al., 2007), first passage times have been used as measures of comprehension (Stewart et al., 2004), transitions as measures of information processing and integration (Pieters et al., 1999; Pieters and Warlop, 1999).

Systematic sequences of fixations and saccades on scenes are called "scanpaths" (Noton and Stark, 1972), and have been studied in advertising (Lohse, 1997; Pieters et al., 1999), web and computer usability (Jacob and Karn, 2003), and choice research (Raaij, 1977; Russo and Leclerc, 1994). Scan paths have either been summarized descriptively or graphically, or through switching matrices (Pieters et al., 1997). If interest is in the extent of parafoveal and peripheral processing, and in the influence of memory and other top-down factors, *distance measures* are useful such as the length of saccades between fixation points (Irwin, 2004) or the angle or horizontal and vertical direction of saccades. While the previous measures emphasize the saccades and their sequence, others emphasizes fixations and their duration.

If interest is in the processing intensity at specific locations in the scene, temporal aggregation across individual fixations is done to arrive at *density and duration measures*, such as fixation frequency and gaze duration (sum of fixation durations or total time spend) (see Inhoff and Radach (1998) for specific measures used in reading and cognitive processing research). These measures are probably the most frequently used measures in eye tracking for academic and applied purposes (Pieters and Wedel, 2004). Because under normal conditions the correlation between fixation frequencies and gaze durations is quite high (Inhoff and Radach, 1998), the specific density or duration measure to use is dictated by the theory under consideration. For instance, Wedel and Pieters (2000) used fixation frequencies on print ads to model information storage and retrieval grounded in theories that predicted a constant information rate per fixation. Pieters and Warlop (1999) found that fixation durations during stimulus-based choice were reduced under time pressure, and that longer fixation durations and increased fixation durations predicted final choice. Aribarg et al. (2008) extend that work by showing how gaze durations arise as a random sum of individual fixation durations on elements of visual scenes. Average fixation durations across all fixations at a location in the scene or at the scene as a whole, have been shown to be related to low level visual factors such as stimulus features (color versus black-and-white), mode (text versus scenes), and visual complexity (Henderson, 2003; Irwin, 2004; Rayner, 1998).

5.2 Space

At the spatially most disaggregate level, information is available about eye fixations at a pixel-by-pixel basis. Because the distribution of fixations is typically not uniform across the scene, the transition matrix between exact fixation locations is sparse, which seriously complicates analyses. Take for example the Mercedes ad in Figure 2.1, where less than 15% of the space (pixels) attract all fixations and the rest receives none. Individual fixation locations are often used in direct graphical representation of eye movements through gaze plots, hot-spot plots, and gaze-replay clips provided by the standard software that comes with most eye-tracking equipment. Such hot-spot plots place a colored window over the target stimulus, with lighter (red, orange, yellow) colors reflecting locations receiving most, and darker colors reflecting locations receiving least eye fixations. Such graphical representations are a useful first step in understanding the data. Figure 5.1 provides an example of a gazeplot, based on a study by Liechty et al. (2003). The graph shows an example of a full-page ad and its editorial counterpage and the eye-movement pattern of one respondent. The fixations are shown as circles, saccades connecting them as lines. Unobserved visual attention states, local (green) in which detailed information is collected, and global (red) in which new interesting locations are identified, inferred by a (Hidden Markov) model from the eye movements are shown as colors. The figure illustrates that the person intermittingly searches the stimulus for new features/objects, which are then investigated in more detail.

Analyzing individual fixation locations at the level of pixels can and has been done (Van der Lans et al., 2008a,b), but it is computationally intense. The variation of fixation points across individuals, that is, the spatial variance of the x–y locations in pixel coordinates, has been used as an effective measure for attention capture and guidance by the visual stimulus, in particular for moving images (d'Ydewalle et al., 1988; Teixeira et al., 2008).

Spatial aggregation beyond the level of the individual fixation is often justified because the visual field during a fixation is larger than the exact x–y coordinate where the fixation lands (Anstis, 1974; Findlay

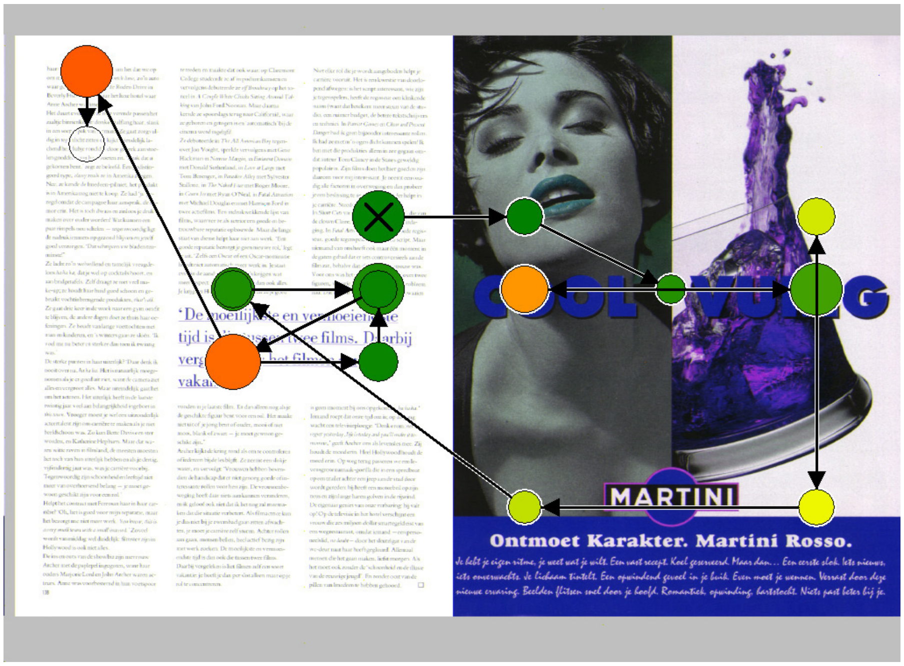

Fig. 5.1 Print ad and counterpage, with eye-fixations as circles and local/global attention states in green/red.

and Gilchrist, 1998; McConkie, 1983), that the eye is never exactly still due to microsaccades, tremor and drift, and that spatial resolution of commercial eye trackers is about 0.5° of visual angle. Aggregation of eye movements can take place across space by superimposing grids, and by singling out meaningful objects or parts in the scene.

Aggregation across a spatial grid is typically applied when the interest is in analyzing sequences of fixations over time and space, when the approximate distances of fixation locations and their spatial organization need to be retained, and there is no *a priori* interest in specific objects in the scene. Spatial grids are particularly useful because of the interest of spatial location and design for marketing purposes (in shelf design, feature advertising, banner advertising, sponsored search, and packaging design, for example). Typically regular grids of 4×4 or up to 6×8 cells have been used (Liechty et al., 2003; Viviani, 1990), but some have used irregular spatial grids that reflect content

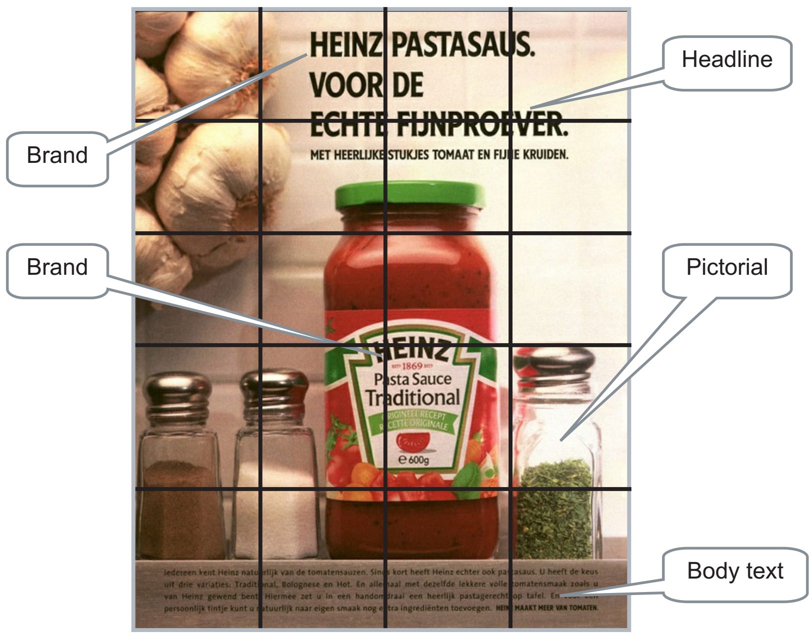

Fig. 5.2 Print ad: Design elements and spatial grid.

(Dreze and Hussherr, 2003). The choice of grid size balances the level of detail against data sparseness, Figure 5.2 shows an example for a print ad. In some cases, the grid cells are content-analyzed by experts to provide ratings of semantic characteristics, such as "informativeness," the number of words of text, or occurrence of people and faces, that can be related to eye movements (Antes, 1974; Buswell, 1935; Henderson and Hollingworth, 1998; Loftus and Mackworth, 1978). This combines grid-based aggregation with the approaches discussed next.

Aggregation of fixations across meaningful objects can focus on substantive or structural objects. Substantive objects are natural or man-made objects with a semantic meaning in the scene, such as Marilyn Monroe, her earrings or beauty mark in the Mercedes ad. Fundamental research on target search and scene perception in psychology and engineering has focused on natural objects that are semantically inconsistent or incongruous in the scene, such as a monkey in a farm scene, or a longdrink in a laboratory setting (e.g., Henderson and

Hollingworth, 1998; Loftus and Mackworth, 1978), which is relevant for research on advertising originality (Pieters et al., 2002). Incongruity has been specifically studied for print ads by Pieters et al. (2008). Structural objects are objects with a syntactical meaning in the scene, i.e., a certain fixed role or function, such as the brand being the target in the ad, and the text and pictorial containing the key message. Marketing research has emphasized structural objects, such as headlines, body text, pictorial and brands in print advertisements (Rosbergen et al., 1997; Treistman and Gregg, 1979), the same elements and price and promotion elements in feature ads (Pieters et al., 2007), brand and SKU facings on shelves (Chandon et al., 2007, 2008; Van der Lans et al., 2008b), warnings in beer and tobacco ads (Fox et al., 1998), and popups and banner ads on webpages (Dreze and Husserr, 2003). Aggregation of eye movements across objects takes place when interest is more on these objects and their relationships, than on their spatial arrangements and distances in the scene, because these latter characteristics are essentially lost in the aggregation. In addition to their semantic and syntactical content, the surface size of objects in scenes is relevant because unlike cells in grid-based aggregation, these objects usually have different sizes, and even if eye fixations would be distributed randomly across the scene, they would exhibit a monotone relation with the surface size of the objects (Yantis, 2000, p. 80). This calls for eye-movement metrics that are independent of the surface size of objects (e.g., fixation frequency or gaze per unit surface size), or for controlling the surface size of the objects in statistical analyses on them (Pieters and Wedel, 2004). Figure 5.2 provides an example of a print ad with structural objects. Where to date the coding of objects in scenes is mostly accomplished manually (see e.g., Wedel et al., 2008), the identification of such objects with computer vision methods (Stockman and Shapiro, 2001) holds promise.

5.3 Individuals

Sample sizes of participants in eye-tracking research have typically been relatively small, and thus individual differences have seldomly been described or modeled explicitly until recently. This is particularly the case in basic cognitive psychology work on reading and scene perception (Rayner, 1998; Rayner and Castellano, 2007) where emphasis has been on stimulus factors influencing eye movements. In one of the first positive examples, Krugman et al. (1994) collected a large dataset on 328 young consumers whose eye movements were recorded during exposure to cigarette ads. The authors reported among others, the percentage of participants fixating on the warnings, and the average time in seconds until the first fixation on the warnings. These measures combine the three aggregation modes, time, space, and individuals. Similar metrics, such as the percentage of participants skipping versus fixating the brand, text or pictorial in ads at least once, as a measure of attention selection have been used (Pieters and Wedel, 2004; Pieters et al., 2007). Measures of variability of fixation points across individuals have been discussed above (Teixeira et al., 2008). Specific features of the visual stimuli are related to the eye movements, aggregated across time, space, and/or individuals.

5.4 Pupil Diameter

The pupil is the opening at the center of the iris of the eye through which light passes. The pupil constricts to about 1.5 mm in bright light conditions and dilates to about 8 or 9 mm in dim light conditions. In addition, pupil diameter has been related to various cognitive and emotional processes, and is thus of interest here. However, developing metrics from the pupil's diameter (also termed pupil dilation or pupillary size) is nontrivial because it constantly and spontaneously fluctuates during waking hours (Loewenfeld, 1999), to optimize incoming amounts of light, and measurements need to take this into account, to which we return later. Task-evoked pupillary responses (TEPR) can be obtained, as the average diameter before (the baseline) and after a specific task or event is introduced, which effectively controls for background fluctuations and individual differences.

5.5 Visual Stimuli

In analyzing eye movements, some studies have coded specific details of the visual image (Van der Lans et al., 2008a,b; Teixeira et al., 2008).

In particular, the RGB color decomposition of the image at the level of individual pixels can be derived with standard imaging software, and from that luminance levels per pixel can be derived. Edges of objects in the image can then be derived from the luminance values (as points in the image where the 2nd derivative of luminance equals zero). See for example Stockman and Shapiro (2001) for details. The extraction and use of such features of the image probably mimicks the way the visual brain (e.g., areas V1, V4) decomposes image information (Marr, 1982). These features can be used to derive textures, visual groupings, and objects from the image. But, to date relatively little progress has been made in utilizing this information in analyzing eye-tracking data, in spite of the fact that in the image processing literature much progress has been made in developing algorithms that facilitate the recognition of such structures in images (Stockman and Shapiro, 2001). Most of the studies to date use manual coding of visual groupings constituting, for example the layout of shelves and brand facings on them (Van der Lans et al., 2008a,b), objects, that is brands or banner ads (Dreze and Hussherr, 2003; Teixeira et al., 2008), or structural visual elements, including the pictorial, text, headline and brand of print (Pieters and Wedel, 2004) and also price and promotion elements of feature ads (Pieters et al., 2007). In the case of moving visual images, such manual coding may be quite time consuming. Teixeira et al. (2008) coded "separation," indicating whether the brand in the video stream is well-separated from its background or not; "cardinality," which captures how many times a brand appears nonconsecutively; and "duration," which indicates how long a brand with the same cardinality was present for every frame in a commercial. These measures can be derived for other objects as well. For dynamic stimuli, the speed at which different scenes are presented is also an important variable, called pacing. Pacing is reflected in discontinuities in the video stream and accomplished by cuts and edits (Germeys and d'Ydewalle, 2007), and were used by Teixeira et al. (2008). The coding of features and elements and objects in images and movies for purposes of eye tracking is still in its infancy, and rather than currently used manual coding, the application of algorithms developed in computer vision holds much promise (Oliva and Torralba, 2001; Stockman and Shapiro, 2001).

Progress has been made with respect to coding properties of the image-as-a-whole. The amount of visual (featural) complexity can be assessed by the file size of the JPG-compressed image in kilobytes (Calvo and Lang, 2004; Sprott et al., 2002). Compression algorithms, such as for the JPG formats cause larger compressions if visual images contain little visual detail, color, contrast, and contain many redundancies, for instance when their surface is homogenous Sprott et al. (2002). This makes file size a suitable general measure of the visual complexity of images. Research has found the file size of images such as charts, web images, and photos to correlate highly and significantly (0.82) with human judgments of visual complexity (Calvo and Lang, 2004; Donderi, 2006). Teixeira et al. (2008) and Pieters et al. (2008) used these compression algorithm measures in eye tracking.

Pieters et al. (2008) focus on the visual complexity of ad images, building on information-theoretic principles derived from Berlyne's (1958, 1971, 1974) theory of visual complexity. There are ten aspects of their measure of visual complexity, including the number, irregularity, heterogeneity, detail, contrast, size, masking, and asymmetry and irregularity of arrangement of objects, as well as the heterogeneity of the background. Although an objective classification scheme is available, manual coding of the images is as yet still required. We refer to that work for a detailed description of the typology.

5.6 Model-Based Inference

5.6.1 Reverse Inference

Eye-movement recording rests on the assumption that patterns of fixations and saccades are overt measures of the covert attention, cognitive and perhaps other processes that are of prime interest. These are described in detail in the next section. Russo argued in 1977 "Eye fixations are information acquisition responses. Rarely are they anything else. This means that using them, to identify cognitive processes is like using only the tip to judge the shape of an entire iceberg. [...] Cognitive processing can be partitioned into two types of activity, the acquisition of information and the operations that are performed on this input. Any strategy for performing a cognitive task, such as consumer decision making, will exhibit a characteristic pattern of information acquisition and processing. The task of the researcher is to identify the consumer's strategy from only what is observable." Research has supported Russo's claim that eye movements reflect information acquisition processes, and has shown that they reflect covert attentional and other cognitive processes in general. This is because covert shifts of visual attention and overt eye movements are tightly coupled (Rizzolatti, 1987; Rizzolatti et al., 1994; Findlay and Gilchrist, 2001), and guided by common cortical mechanisms (Corbetta and Shulman, 2002; Fuster, 2003; McFadden and Wallman, 2001). Thus, although overt and covert attention can be dissociated in laboratory settings, under normal conditions the eyes follow attention and attention follows the eyes. Henderson and Hollingworth (1998) use the metaphor of a rubber-band to describe this coupling.

The goal of eye-tracking studies in marketing then is to understand attention processes to marketing stimuli in real-life settings, and their determinants and consequences. This may be challenging. As a direct consequence of the lack of experimental control in such natural exposure conditions, the large data streams from eye-movement recordings, the stochastic nature of eye-movement data, and the heterogeneity across stimuli and participants, statistical models have become critical tools to gain insight in the underlying unobserved (latent) attention processes from eye-tracking data, post-hoc, which Feng (2004) calls "reverse inference." The statistical tools that we discuss here are grounded in theory — albeit to different degrees — and provide a basis for testing theories of attention and their further development. Other methods have been or can be used to describe eye movements, including decision trees (Chandon et al., 2007). But, in spite of their success in describing overt eye-movement patterns, these methods are not directly grounded in theory and therefore not discussed in detail here.

5.6.2 Bayesian Modeling

The Bayesian statistical framework provides a formalized way of learning about the parameters of a model from data that seamlessly addresses the challenges of analyzing eye-movement data. Uncertainty about the parameters is represented through a prior distribution and inference proceeds conditional upon the data obtained. Bayes' theorem shows how the prior distribution of the parameters can be updated with the information on the parameters contained in the data (through the likelihood), which results in a posterior distribution for the parameters. The interpretation from a subjectivist point of view is that it measures the personal uncertainty about the parameters characterizing the attention process, after the data has been observed. The posterior distribution can take complex forms for realistic models. A breakthrough in the application of the Bayesian paradigm was the realization that in many cases the expressions for the posterior distributions of the parameters can be factored into simpler expressions that can be recursively sampled from, using so called Markov Chain Monte Carlo methods (MCMC, Gelman et al., 1995; Geman and Geman, 1984). The Bayesian paradigm has gained great popularity in statistics and its applications, because it allows a flexible and robust framework for developing and estimating statistical models that facilitate realistic description of psychological data — and eye-movement data in particular, as it allows for inference in finite samples and effortlessly accommodates random coefficients, hierarchical effects, missing data, and latent variables. In addition, the subjective interpretation of prior and posterior probabilities and its consequences for theory testing is appealing from a theoretical perspective and as a consequence is gaining ground.

Importantly, the Bayesian paradigm has recently been recognized not only as an integrated statistical framework for inference, testing and decision making, but also as a theoretical framework for understanding visual perception in general, including such processes as top-down effects in high level vision (Kersten, 1999), feedback mechanisms in the visual cortex (Lee and Mumford, 2003), object recognition (Kersten et al., 2004), and optimal scan paths in visual search (Najemnik and Geisler, 2005). Research in those areas shows that the human brain updates prior information with new evidence obtained from visual data using the theorem proposed by Thomas Bayes. We therefore advocate the use of Bayesian methods for the construction of realistic models to infer cognitive processes from eye-movement data (see Aribarg et al. (2008), Van der Lans et al. (2008a), Pieters et al. (2002), Wedel and Pieters (2000)).

5.6.3 Models at Various Levels of Data Aggregation

Although an extensive discussion of the statistical models developed for eye-movement data is beyond the present scope, we provide a few key models and model components. These developments are closely tied to the dimensions of the eye-tracking data — time, space, stimuli and individuals — that we described in Section 5.5. Typical disaggregate eye-movement data are of the form $(x_{i,t,s}, y_{i,t,s}, d_{i,t,s})$, for $i = 1,\ldots,N$ individuals, $t = 1,\ldots,T$ time points (fixations), $s = 1,\ldots,S$ stimuli and x and y the coordinates of the location of the tth fixation, and d its duration. Often, R regions of interest (ROI) are defined, $r = 1,\ldots,R$ as described in Section 5.2 through a mapping of the image coordinates $(r_{i,t,s} : x_{i,t,s}, y_{i,t,s} \in \Re_r)$. If fixations are retained at the level of the ROIs, the data becomes $(r_{i,t,s}, d_{i,t,r,s})$, with $\Re_r$ denoting ROI r. If data is aggregated over time, one obtains $d_{i,r,s} = \sum_t d_{i,t,r,s}$ the gaze duration of individual i on ROI r of stimulus s, which if aggregated across individuals becomes $d_{r,s} = \sum_i d_{i,r,s}/N$. Aggregation across one or more of the layers of the person-time-space cube offers different insights into attention processes.

Eye movements aggregated across individuals, time, and space. Aggregation of eye-tracking data in time, space or across individuals makes it more difficult to incorporate or test specific micro-level theories of visual perception or cognition. Such aggregation requires assumptions of the equivalence of macro (ad-level) with micro (time–space–person level) attention processes, or different macro-level models. Pieters and Wedel (2004) and Pieters et al. (2007) analyze advertisements-by-ROI data matrices, in which gaze durations are aggregated both over time, space, and individuals. This high level of aggregation was necessary because of very large samples of thousands of individuals and hundreds of ads. These authors use regression models of average gaze durations on the ROI's (brand, text, pictorial, etc.), in which the dependent variable is log-transformed to accommodate its positive support and right-skewed distribution and they incorporate theories of attention capture and transfer, and target distinctiveness and distracter heterogeneity into these models through the specification of independent variables. These models are of the type: $\ln(d_{r,s}) = x'_{r,s}\beta + \varepsilon_{r,s}$, with $x'_{r,s}$ a vector of explanatory variables chosen to reflect specific theoretical concepts or predictions.

Eye movements aggregated across time and space. For individual-level gaze durations on spatial or content-based ROIs of ads or other visual stimuli, gamma and log-normal duration regression models are suitable (Rosbergen et al., 1997; Pieters and Wedel, 2004). If instead fixation frequency counts are retained at the level of ROIs, then count data regression models, such as the Poisson and Negative Binomial Regression (NBD) models can be applied (Pieters et al., 2002; Wedel and Pieters, 2000). These models deal with the stochastic nature of the individual-level gaze data and its (positive) support. Random Stopped Sum (RSS) models provide an integrated descriptions of gaze durations as the sum of the fixation durations over the number of fixations on an ROI (Aribarg et al., 2008). These RSS models generalize models for gaze durations and fixation frequencies and provide a more principled account of fixations on areas of interest. In these models, typically the surface sizes of design elements of ads have been included as independent variables to predict gaze. These models are of the type $g^{-1}(E[d_{i,r,s}]) = x'_{r,s}\beta$; $d_{i,r,s} \sim f$, with $g^{-1}(\cdot)$ a link function such as the natural logarithm, $E[\cdot]$ the expectation operator, and f the chosen distributional form.

Individual-level eye movements with heterogeneity. In many cases, generalized linear regression type-models such as the ones described above for individual-level data need to include a hierarchical structure to describe heterogeneity in eye-movement patterns across ROIs. This is done through the specification of a statistical distribution of the model parameters across the sample of individuals capturing the attentional process in question. Discrete heterogeneity distributions (the Multinomial) have been used (Rosbergen et al., 1997) giving rise to finite mixture or “latent class” formulations that offer advantages of interpretation of groups of individuals as segments with different eye-movement patterns. More recently, continuous distributions have been favored that allow for individual level inference. While Wedel and Pieters (2000) and Pieters et al. (1999) use the mathematically convenient conjugate gamma specification of heterogeneity in their Poisson fixation frequency models leading to the NBD, Aribarg et al. (2008) specify the more flexible multivariate normal distribution of effect parameters across individuals, which, however, somewhat complicates the estimation algorithm. These models can be written as $g^{-1}(E[d_{i,r,s}]) = x'_{r,s}\beta_i$; $d_{i,r,s} \sim f$; with $\beta_i \sim N(\mu_\beta, \Sigma_\beta)$ normally distributed with mean vector and covariance matrix as specified, or with $\beta_i \sim MN(\pi)$, multinomial with π a K-vector of class probabilities.

In extensions of these continuous distribution hierarchical models, the design matrix describing experimentally varied task conditions are introduced as independent variables into the prior distribution for the parameters. This allows one to model top-down effects on attention measures by experimentally separating them from bottom-up effects (Pieters and Wedel, 2007a,b; Van der Lans et al., 2008a,b). Modeling the effects of these experimental manipulations through the prior is consistent with theoretical frameworks for top-down effects in visual perception (Kersten, 1999; Lee and Mumford, 2003). If these experimental conditions are contained in the design matrix z_i, for individual i, then the model is

$$g^{-1}(E[d_{i,r,s}]) = x'_{r,s}\beta_i;\ d_{i,r,s} \sim f; \beta_i \sim N(z_i'\phi, \Sigma_\beta)$$

Although methods have been developed that allow one to accommodate individual-level heterogeneity in data aggregated across individuals (Chen and Yang, 2007; Musalem et al., 2008), these methods have not found their way into eye-tracking data as yet. Since disaggregate data is often available but sometimes massive, the application of these methods would be motivated by trading off storage for computational costs.

Eye-movement sequences in time. To describe and understand sequences of eye movements over time, Markov models can be used and have been shown to reveal powerful insights (Harris, 1993; Pieters et al., 1999; Stark and Ellis, 1981). These models describe successive fixations on ROIs as: $P(r_{i,t,s}|r'_{i,t-1,s}) = \pi_{r',s}$ an R-vector of probabilities that sum to 1. Explanatory variables, such as surface sizes of the

ROIs can be introduced in the probabilities through a logit-link. For instance, Pieters et al. (1999) developed a Markov model to describe the sequence of eye movements (saccades) between and within ROI's in print advertisements. Heterogeneity was included through a conjugate distribution, which simplified estimation. Consumers were repeatedly exposed to the ads. Pieters et al. (1999) found that, although the overall fixation frequency (marginal distribution of the switching matrices) decreased across repetitions (people looked less), the patterns of switching between objects in the ads remained stable, and obeyed a quasi-symmetric Markov chain. This provided empirical evidence for Noton and Stark (1971) scan path theory.

Hidden Markov Models (HMM) have proven to be particularly powerful methods to capture time-paths of unobserved attention states and have become a central tool in a stream of work on eye movements (Van der Lans et al., 2008a,b; Liechty et al., 2003; Rimey and Brown, 1991; Wedel et al., 2008). These HMMs conceptually extend finite mixture models. Whereas the latter models are based on the assumption that a number of homogeneous groups or states underlies an eye-movement sequence, the HMM allows consumers to switch between these states over time, according to a first-order Markov process. The probability of a sequence of fixations of a consumer i, $V_{i,s} = \{r_{i,t,s}; t = 1, \ldots, T\}$, is written as a HMM $\Pr(V_{i,s}) = \sum_{w_1=2}^{W} \cdots \sum_{w_T}^{W} \prod_{t=2}^{T} \pi_{w_t|w_{t-1}} \Pr(V_{i,s}|w_t)$, with $w_t = 1, \ldots, W$ hidden states, $\pi_{w_t|w_{t-1}}$ transition probabilities between the hidden states and $\Pr(V_{i,s}|w_t)$ the probability of observing the sequence of eye movements given a hidden state. These could themselves obey a Markov model, as described above. For example, Wedel et al. (2008), developed a HMM to describe eye movements on spatial ROIs on ads. The observed eye movements on a spatial grid overlaid on ads were described with a Markov model, while it was assumed that two hidden states underlie the observed switches between cells of the grid, with in one (local) state switches between neighboring cells more likely, and in the second (global) state switches between non-neighboring cells more likely.

The assumption of underlying discrete states in HMMs is conceptually appealing, because it conforms with behavioral theory on mental states. Nevertheless, whereas HMMs capture dynamics in eye movements through such discrete underlying states, an alternative formulation is through dynamic (generalized linear) models that allow parameters to evolve over time continuously, for example through a first-order autoregressive process. These models can be used to capture the effects of dynamic stimuli such as TV commercials where exposure to the stimulus is expected to create a context that changes the parameters of the attentional process itself (Teixeira et al., 2008).

Individual eye-movement sequences in space and time. The analysis of fixation sequences at the most disaggregate level, yielding individuals by time by location data, is particularly challenging, not in the least because of the massive volume of data, and the need to characterize visual input obtained at the level of x–y fixation locations coded in pixel coordinates. Very little work has as yet addressed these challenges, but spatial poisson processes have been used to describe fixation patterns on images at the pixel-level (Van der Lans et al., 2008a,b). The models in question are highly tailored to the analysis of eye-movement data and model the landing location of each fixation as a choice between any one of the more than 1,000,000 pixels on a visual display.

The above models are grounded in theories of visual attention and eye movements, the extent to which that is accomplished depending on the nature of the problem, the availability of theories, the aggregation level of the data and computational limits of the statistical models. The approaches developed capitalize on the Bayesian frameworks for describing perception and for statistical inference. The modularity of Bayesian models allows one to combine several of these models (for example, count data regression, hierarchical and hidden Markov models) as model-components into larger modeling frameworks, the compilation of these components being dictated by the nature of the substantive problem and the theory predicting the attentional phenomenon in question.

6

A Theory of Attention to Visual Marketing

The amount of information that is transmitted through the optic nerve, consisting of circa 1.5 million axons is estimated to be in the range of 108 $\sim$ 109 bits per second (Koch, 2004), and far exceeds what the brain can process. Therefore, the brain needs to rely on mechanisms that select a subset of relevant information for enhanced processing. This process of selection and focalization is called attention, first defined by James in 1890. Attention is a limited capacity, selective process that serves to coordinate the perception-action cycle and to maintain goals over time (LaBerge, 1995). Visual attention is the psychological construct of key interest in much eye-movement research in marketing (Pieters and Wedel, 2004). When a particular location or object in a scene is selected through attention, processing of it is enhanced, and processing of nonselected locations and objects is simultaneously suppressed. This is accomplished through bottom-up and top-down mechanisms. The first evolves from the visual environment itself, i.e., the visual stimuli; the second mechanism is driven by the behavioral relevance of aspects of the visual scene and originates from higher cortical centers, i.e., from goals, memory and states and traits of the individual.

Figure 6.1 summarizes the main structure of the theory of visual attention that is described next.

The theory of attention to visual marketing describes the determinants of attention to visual marketing stimuli, as reflected in eye movements, and the influence of attention on communication effects that are of direct interest to marketing, such as memory, preference, and choice. The theory distinguishes top-down sources of influence on attention, emanating from prior states and traits of consumers, such as their expectations, goals and emotions, and bottom-up sources of influence of attention, emanating from the features of the stimuli that consumers are exposed to. The two sources of influence combine to determine covert attention patterns as revealed in overt eye movements. The theory distinguishes the *informativeness* of stimuli and their specific locations and objects, which derives from their goal-relevance (top-down), from the *salience* of stimuli and their specific locations and objects, which derives from perceptual feature contrasts (bottom-up). The theory specifies that salience and informativeness combine (additively or multiplicatively) to affect attention and other processes. Thus, strongly salient stimuli will capture attention even if consumers are not particularly sensitized to them (as in Yantis and Jonides (1990)). Conversely, consumers will vary in their attention to stimuli and find them differentially informative depending on their specific goals and expectations (as in Öhman et al. (2001)), almost independent of the specific stimulus configuration. Yet, attention is greatest when salient stimuli are informative, and the other way around (as in Folk et al. (1992)). The theory reconciles divergent views of the determinants of attention, systematically accounts for informativeness and salience, and relates these to relevant visual marketing effects. Let us now examine it in more detail.

As described in Section 3, upon exposure to a visual scene, such as an advertisement, website or bilboard, basic perceptual features including colors, luminance and edges are extracted automatically and in parallel (Treisman and Gelade, 1980; Wolfe and Horowitz, 2004). They are used to rapidly and pre-attentively segment the image (Duncan and Humphreys, 1992; Wertheimer, 2001). Here, broad organizational principles such as horizontal and vertical dimensions play a role (Oliva and

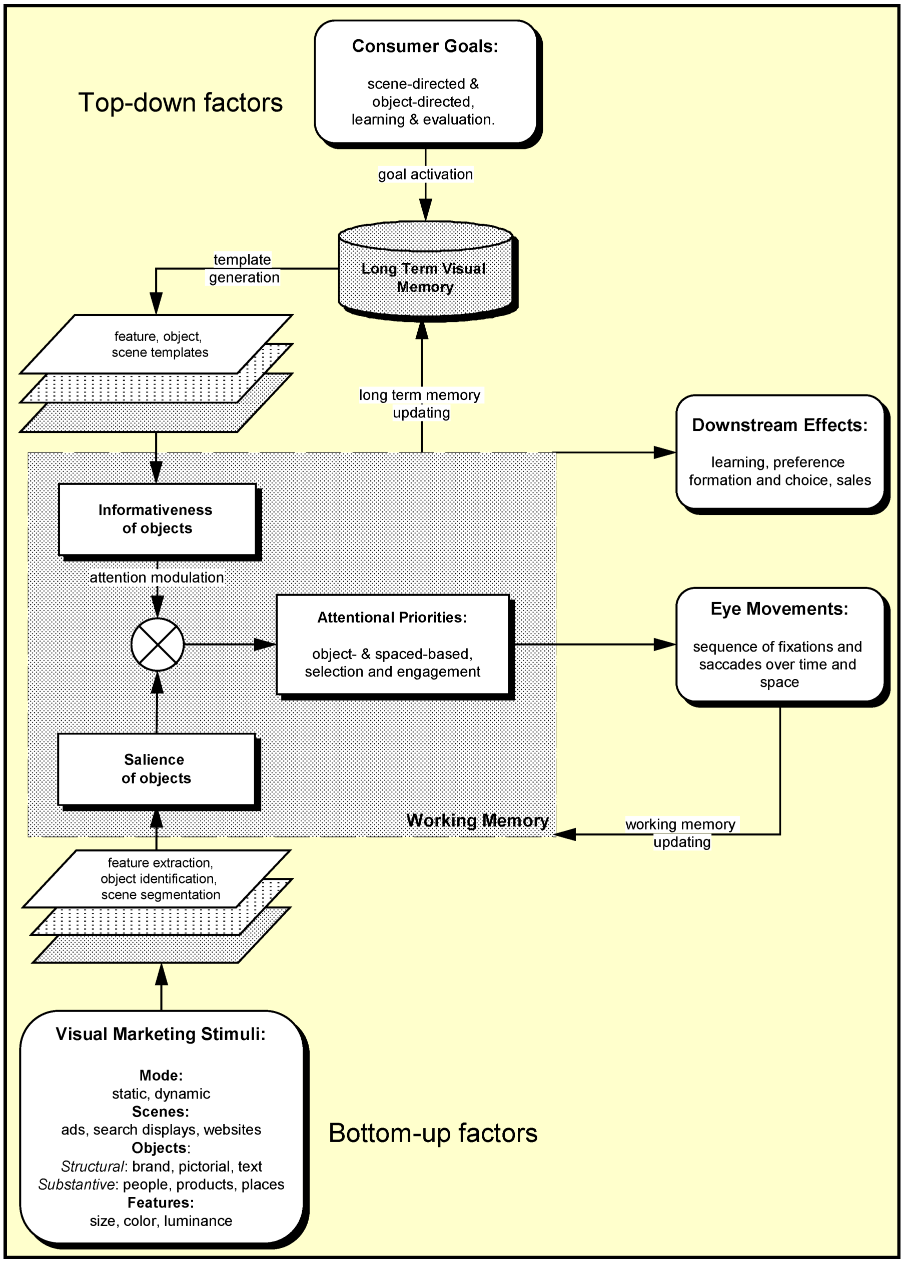

Fig. 6.1 A theory of attention to visual marketing.

Torralba, 2001), as indicated in the lower half of Figure 6.1. The horizontal layout of product shelves in supermarkets is an example. Rapid segmentation enables people to grasp the "gist" of the scene (Friedman, 1979), such as whether it is outdoor or indoor, natural or man-made, a beach or forest scene and so forth, perhaps even before specific objects or features within it are actually identified, and the meaning of the scene itself strongly guides the interpretation of objects. That is, the gist of the scene, such as that of a convenience store, immediately prime schemas of interrelated objects with a high likelihood of being present — candybars, softdrinks, potato chips — and suppresses objects unlikely to be present — stereo sets, CD players, and cars. Thus the scene's gist can often be apprehended before specific objects inhabiting the scene: the whole comes before the parts. This process is thought to be driven primarily by low spatial frequency information in the visual scene (Oliva and Torralba, 2001).

Upon exposure to the scene, the salience or conspicuity of specific locations and objects in the scene is pre-attentively determined, based on their feature composition. Salience is not simply determined by the intensity of the individual features of an item or location. Rather, the extent to which an item or location contrasts on basic perceptual features with its surroundings (e.g., luminance, color, orientation, texture; Wolfe and Horowitz, 2004) determines the extent to which it is visually conspicuous. Conspicuity is represented in a salience map (see Section 3). The salience map is a topographic map coding the visual significance of all locations in a scene (Koch and Ullman, 1985), and it guides subsequent focal attention. It provides an efficient mechanism for attention guidance, since it enables shifting the focus of attention to locations in order of decreasing salience, through the interplay of winner-takes-all mechanisms (WTA: with multiple locations competing for attention the most salient one gains it) and inhibition of return mechanisms (IOR: the previously attended location is suppressed to prevent attention from oscillating between locations of high salience) (Itti and Koch, 2001). This leads to relatively stable scan paths of eye movements across stationary scenes (Noton and Stark, 1971; Pieters et al., 1999). The salience map is build-up from local contrast in basic perceptual features (Wolfe and Horowitz, 2004). A weighted combination of these features forms the salience map. Much is known in basic attention research about the influence of such stimulus-based (bottom-up) salience on attention (Itti and Koch, 2001; Parkhurst et al., 2002).

An object "pops-out" and is found instantaneously on the first eye fixation based on pre-attentive processes, when it is dissimilar from all distractors on a single perceptual feature and when all distractors are similar on that feature (Duncan and Humphreys, 1992), such as when searching for a Heinz green ketchup bottle among a set of homogeneously red distractor brands. Such bottom-up control of attention is mostly involuntary. Less is known about the influence of top-down, voluntary mechanism in guiding attention, which are represented in the top part of Figure 6.1.

Yet, both bottom-up and top-down processes influence attention. That is, to the extent that an object has perceptual features that are deemed informative by consumers in the pursuit of their current goals it will attract more attention. Top-down factors such as consumers' search goals ("find the new Heinz bottle," or "...the cheapest cakemix"), memory for perceptual features and objects ("what is the color of the new Heinz bottle?") (Treue, 2003). This takes place by selectively enhancing visual features that are diagnostic and by selectively suppressing features that are nondiagnostic (Lee and Mumford, 2003; Pieters and Wedel, 2007a). Such voluntary, top-down enhancement and suppression of perceptual features is effortful, usually slower and typically limited to only one or two features (Vogel et al., 2001; Wolfe et al., 1990). Such attention modulation is indicated in the center-left part of Figure 6.1, taking place in working memory. Thus, perceptual features inhabiting spaces in scenes are pre-attentively located and bound into objects by focused, slow attention (Treisman and Gelade, 1980), due to top-down influence.

The preceding account is predominantly spatial and consistent with theories of early selection (Broadbent, 1958). In addition, it has been shown that attention serves to distinguish relevant objects in a scene by giving early priority to the processing of features (including spatial location) that define the object, presumably through synchronous firing of (populations of) neurons in the visual brain (Singer and Gray, 1995). Space- and object-based accounts of attention are complementary, since both spatial uncertainty (where is it?) and object identity uncertainty (what is it?) may exist during natural exposure of individuals to complex scenes, such as ads or shelves (Van der Lans et al., 2008a,b). Spatial and identity uncertainty require qualitatively distinct processes (Section 3; Koch, 2004; Ungerleider and Mishkin, 1982). Salience-based attention processes, reflected in pop-out, serve as circuit breakers to top-down processes and guide attention to novel and unexpected stimuli.

Multiple objects that are simultaneously present compete for neural representation and thus attention (Desimone and Duncan, 1995; Duncan, 1980, 1996; Lamberts, 2000; Tsotsos, 1990). Tjan (1999) proposes that memory modules for object recognition are attached to a single visual-processing pathway. Object-identification occurs independently at each site and its response latency and inaccuracy are monotonic-decreasing function of the extent of the match with the incoming visual signal. The observer's response — recognition of the object — is the first-arriving response from any of the memory sites (Lamberts, 2000). Competition among objects is based on their perceptual salience, on local feature contrasts, and on their goal relevance or informativeness in the scene. Biased competition due to goal relevance is achieved through top-down modulation (consistent with theories of late selection, Deutsch and Deutsch (1963)). It allows the visual system to enhance the processing of informative (i.e., goal relevant) stimuli and suppress uninformative stimuli. This implies that the informativeness of objects does not reside in the scene itself, and that thus it is not the case that some stimuli are "intrinsically" more informative than others, but resides in the goals of the person looking at the scene (Yarbus, 1967). Once triggered, a goal prime templates or prototypes in memory (Friedman, 1979; Tipper and Weaver, 1998). Such memory templates need not be one-to-one representations of ads, products, brands or whole scenes, but can be fairly flexible constellations of features (Coca Cola is red), objects (McDonalds is Golden Arches and Ronald), scenes or dimensions of them (car ads are set outdoors, skincare ads are indoors). These templates are matched with the objects appearing in the scene, which happens faster and more accurately if they correspond better to their internal representations (Tjan, 1999; Deco and Zihl, 2001). Thus, objects and locations compete for attention, contingent on control by the current goals of consumers (Folk et al., 1992).

As the theory presented in Figure 6.1 indicates, objects in scenes are prioritized for attention as a result of integrating their salience and their informativeness for the current goal, which have similar effects on attention (Navalpakkam and Itti, 2005; Spratling and Johnson, 2004). Attention selects and engages in these prioritized objects through enhancement and inhibition, reflected in eye movements. Information extracted during fixations updates the attentional priorities held in working memory, and the knowledge in long-term memory, which eventually contributes to downstream marketing effects of interest such as learning (memory), preference formation, and choice.

The theory of visual attention summarized here, emphasizes that as a direct result of selective attention mechanisms being under top-down as well as bottom-up control, our representation of the visual world is dominated by the relevance of the information for our current goals, rather than being aimed at an accurate and complete description. Thus, the informativeness or relevance of visual marketing stimuli varies across these consumer goals. The theory is summarized in Figure 6.1, and may serve as the basis to structure past and propose new research on visual attention in marketing.

7

Insights for Visual Marketing

Visual attention research in marketing has come a long way in the over 80 years since the days that Nixon (1924) painstakingly observed eye movements of consumers paging through a magazine with print ads, while hiding himself behind a curtain in a box. Eye-movement data of large samples of consumers attending to large numbers of ads using infrared eye tracking are currently being produced on an industrial scale, and used by companies to optimize decisions on the design of ads, packages, shelves, billboards, web-pages, catalogs, print, and TV ads. The emerging recognition of the central role of attention in consumers cognitive processes and behavior and the realization that eye movements are an accurate reflection of these processes has furthered the use of eye-movement research in academia as well. The application of theories and models of visual attention, reviewed in this survey, is likely to render such research and the decisions based on it increasingly effective in the future. The availability of metrics and models of attention enables tests of long standing beliefs in visual marketing practice and improvements in theory, the validity of findings, predictions, and recommendations. What are the findings thus far from eye-tracking research in marketing? We will summarize and discuss the findings, using the theoretical model in Figure 6.1 as a guide. Research in marketing has extensively studied visual attention, using eye movements as indicators of it, and the effects of factors residing in the individual (top-down), respectively the stimulus (bottom-up) as antecedents of visual attention. The distinction between top-down and bottom-up effects of attention is a fundamental distinction, not only from the point of view of the theory of visual attention shown in Figure 6.1, but also from a substantive marketing perspective, because these two factors require very different marketing strategies and tactics. Whereas bottom-up effects are generally attained through manipulation of the visual stimulus' design, top-down factors are influenced by targeting segments, and affecting peoples memory, attitudes, goals, and intentions. The results of studies of these effects are summarized in Tables 7.1 (Bottom-up effects, Section 7.1) and 7.2 (Top-down effects, Section 7.2). In addition, we will describe in a separate section (Section 7.3) what is known about determinants and implications of differences in pupil diameter. An important question is whether these determinants of attention ultimately have downstream effects, and we discuss research into these effects in Section 7.4. Figure 7.1 summarizes the framework for the review in this section.

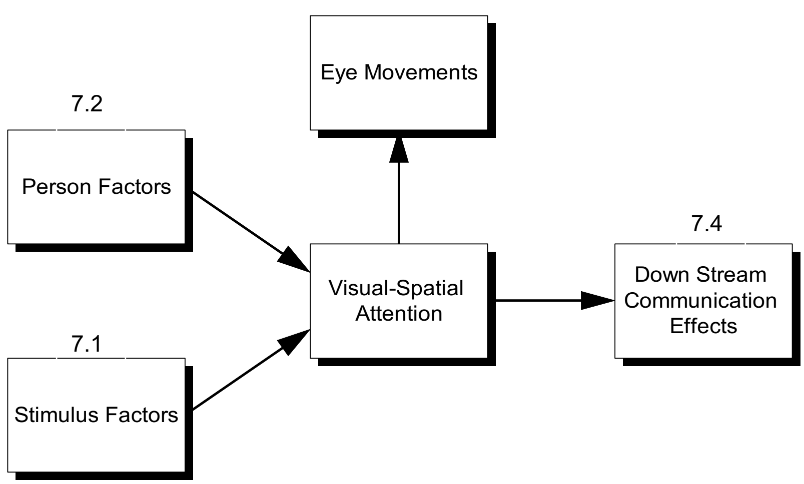

Fig. 7.1 Framework for the review of eye-tracking studies.

Table 7.1 Summary of bottom-up effects.

	Bottom-up effects						
Authors	Size	Color edges	Space	Movement	Image	Competition	Stimulus
Aribarg et al. (2008)	+					+	Print ads
Brasel and Gips (2007)			+	+			TV commercials
Chandon (2002)			+				Shelves
Chandon et al. (2007)			+				Shelves
Dreze and Hussherr (2003)			+				Web-pages
d'Ydewalle et al. (1998)				+			Billboards
d'Ydewalle and Tamsin (1993)			+				Billboards
Fletcher et al. (1995)						+	Health warnings
Fox et al. (1998)						+	Nutrition labels
Goldberg et al. (1999)	+	+	+				Nutrition labels
Janiszewski (1998)	+					+	Print ads
Janiszewski and Warlop (1993)							TV commercials, shelves
Kroeber-Riel (1979)	+						Print ads
Krugman et al. (1994)		+					Heath warnings
Van der Lans et al. (2008a)		+	+				Shelves
Van der Lans et al. (2008b)		+	+				Shelves
Laughery et al. (1993)		+					Print ads
Leven (1991)			+				Print ads
Liechty et al. (2003)		+	+				Print ads
Lohse (1997)		+					Yellow page ads
Lohse and Wu (2003)		+					Yellow page ads
Nixon (1924)	+	+					Print ads

(*Continued*)

Table 7.1 (*Continued*)

	Bottom-up effects						
Authors	Size	Color edges	Space	Movement	Image	Competition	Stimulus
Poffenberger (1925)	+						Print Ads
Pieters et al. (2002)	+				+		Print ads
Pieters et al. (1999)	+		+				Print ads
Pieters and Warlop (1999)	+						Shelves
Pieters and Wedel (2004)	+						Print ads
Pieters and Wedel (2007a)	+						Print ads
Pieters et al. (2007)	+						Feature ads
Pieters et al. (2008)	+				+		Print ads
Rayner and Castellano (2007)	+						Print ads
Rayner et al. (2001)	+						Print ads
Rayner et al. (2008)	+						Print ads
Rosbergen et al. (1997)	+	+	+				Print ads
Russo and Leclerc (1994)							Shelves
Stewart et al. (2004)			+				Brands
Teixeira et al. (2008)	+		+	+	+		TV commercials
Treistman and Gregg (1979)	+						Print ads
Wedel and Pieters (2000)	+					+	Print ads
Wedel et al. (2008)	+		+				Print ads
Zhang et al. (2008)	+	+	+				Feature ads

7.1 Stimulus (Bottom-Up) Effects on Attention

7.1.1 Size

Much is known on the effects of surface size, a bottom-up feature of intermediate complexity. Size clearly matters in capturing attention, but in ways quite different from what has been commonly assumed. Several measures of visual attention, including selection of an object, ad-element or ad, fixation frequency and total gaze duration, are strongly associated with the surface size of print ads, items on shelves, product labels and other visual stimuli, in virtually every study reviewed. Janiszewski (1998) first demonstrated the size-attention effect for print ads. Gaze duration on a print ad as a whole increases by 0.81% for every 1% increase in size (Pieters and Wedel, 2004), for feature ads duration increases by a much smaller amount: 0.22% (Pieters et al., 2007). Zhang et al. (2008) show that surface size has an indirect effect on sales, an effect that is fully mediated by attention to the ad. Differences between these two types of ads may be caused by higher competitive clutter for feature ads, where usually multiple ads occur on a single (newspaper) page. The increase in attention to the text in print ads is almost proportional: a 1% increase in surface size of the text leads to a 0.9% increase in gaze for it (Pieters and Wedel, 2004), presumably due to reading time, which appears almost proportional to text-space. Attention to the brand and pictorial elements increase with increasing surface size at a lesser rate, a 1% increase leading to a 0.3% increase in gaze, i.e., with somewhat less than the square root of the surface size as has been suggested in the early studies by Nixon (1924) and Poffenberger (1925).

Whereas for feature ads the sizes of almost all elements (text, pictorial, brand, price, promotion) increase attention (selection and duration) to the ad, among all print ad design elements only increasing the surface size of the *text element* increases attention to the ad as a whole (Pieters and Wedel, 2004). This is likely due to reading and lead Rayner et al. (2001) to conclude that comprehension of ads is largely based on textual information, and that text may be more important in ad design than previously thought. Pieters et al. (2007) found that for feature advertisements, the text also contributes positively to attention to the ad as a whole, but relative to an optimal composition of feature ads, the text should be decreased in size with about 20%. From the studies reviewed it is not clear to what extent the surface size effect for text is associated with an increase in font-size, or an increase in the number of words. If the latter would be the case, then the interpretation of Rayner et al. may hold and advertisers aiming to maximize attention to the entire ad should consider devoting more space to text. However, most of the studies on print ads to date have failed to find a significant transfer of attention from the text element to the other elements in the ad (e.g., Pieters and Wedel, 2004; Rayner et al., 2001; Witt, 1977). Thus, it seems that attention more readily transfers from the pictorial to the text than the other way around. And although their findings are based on a single study with a relatively homogeneous sample of print ads, Wedel and Pieters (2000) did not find that text significantly contributed to their measures of brand recognition. These findings cast some doubt on the central role of text in print ads for subsequent brand memory, although it may be important for comprehension. More research needs to be done to disentangle the effects of surface size and other low-level features determining salience, from cognitive processing of semantic meaning, or "informativeness." The findings by Stewart et al. (2003), although not pertaining to advertising, reveal that attention to text may have to do with semantic incongruency or complexity. Pieters et al. (2008), however, do not find an effect of message informativeness, incongruency or complexity on gaze duration on the text in print ads. Clearly to the role of the bodytext in advertising needs more emphasis in research. The literature on attention in the context of reading (e.g., Rayner, 1998) is helpful but necessarily limited, because advertisements are usually multi-mode messages with pictorials and text, with the latter having pictorial qualities as well (various colors, fonts, positions), whereas reading research has predominantly emphasized standard text corpus.

The role of the *headline* deserves more study as well. Generally, the idea in advertising theory is that the headline receives attention early, and needs to draw consumers into spending more time with the ad. Few studies, however, have looked into this attention capture and transfer from the headline to the other elements. The headline stands out due to its size, large font type and central position in the ad, summarizes the ad's main message and anchors the meanings in the pictorial. Thus it is often simultaneously highly salient and informative. Rosbergen et al. (1997) investigated attention to the headline, and found it to receive more attention than the body text, while, contrary to the body text, it received attention in each of their consumer segments, which underlines its potentially different role in the ad. Leven's (1991) study showed that scan paths across his ads tended to quickly go to the upper left corner, where usually the headline was located, and where people start reading. Krugman et al.'s (1994) experiment on health warnings also sheds light on this issue. Newly developed warnings had fewer words than the old ones and had larger and bold font, similar to headlines. Effects of salience and informativeness cannot be strictly separated in this study, but the new warnings attracted attention quicker and received longer gazes. Thus, future research may seek to establish the different attention capture and transfer effects in the headline, as compared to the body text, and what the optimal surface and font sizes of headline versus bodytext are. The findings may also be relevant for web design, as Poynter Institute studies on web usability indicate.

The *pictorial*, in most of the studies reviewed, has a high tendency to attract attention, regardless of its surface size. However, the effects of increasing the surface size of the pictorial on attention to it are almost absent, and it has no effect on attention to the ad as a whole (Pieters and Wedel, 2004). Although effects on (indirect) memory have been demonstrated, these are modest at best (Wedel and Pieters, 2000). Based on these current findings, advertisers and agencies would be ill advised to maximize the surface size of the pictorial, regardless of its content, in an effort to maximize attention to the entire ad. However, for feature ads the effect of the size of the pictorial element was found to be larger than for print ads. Pieters et al. (2007) showed that pictorial surface size has the largest effect on attention to a feature ad on an ad-display page, among all elements of the ad. One possible explanation for this is that compared to print ads the competitive clutter for feature ads is much higher, because multiple ads appear on a single page. In conditions of high competitive clutter, the pictorial may be more effective and attract more attention when it is larger. Interestingly though, Pieters et al. (2007) found that in current feature advertisements, the pictorial is still too large. Relative to the optimal feature ad composition that would attract maximum attention to the ad-display page, the pictorial could be decreased in size with close to 40%. Many aspects of the pictorial other than its surface size are not yet very well understood. In particular, aspects of the content of the pictorial, such as the presence of specific objects, their complexity and informativeness are underresearched. One of the first studies that has addressed that issue is the original work by Witt (1977), which showed that higher levels of arousal, primed by the pictorial, promote information intake. Pieters et al. (2008) recently found that visual complexity and informativeness do increase attention to pictorial.

Although few studies report specifically on attention to the *brand or logo* elements, results to date (Pieters et al., 2007; Pieters and Wedel, 2004; Wedel and Pieters, 2000) point to a crucial role of the brand in integrating the information in the ad. While the surface size effect of the brand is similar to that of the pictorial for print ads, attention captured by the brand element transfers more readily to the pictorial and text, than vice versa. Thus, the brand appears to play a key role in routing attention through ads. In addition, the effects of attention to the brand carry over more strongly than those of any of the other design elements to subsequent memory for the brand (Wedel and Pieters, 2000). Thus, ads that succeed in capturing attention to the brand by increasing its surface size, are also likely to succeed in relocating attention to the message of the ads, as contained in the pictorial and text elements, and improving memory for the brand. The results for feature advertising (Pieters et al., 2007) are consistent: the effect of the size of the brand element on attention (selection and engagement) is strong and positive, and for optimal feature ads the size of this element should increase by about 75% over current values used in practice. The brand element itself is deserving of more detailed investigation. It consists of a mixture of pictorial and textual information, which are not necessarily spatially contiguous. The integration of information from these different modes (Rayner et al., 2001) as well as aspects of the spatial distribution of attention are interesting yet challenging topics. The effects of the brand in TV commercials is different from that in print and feature ads (Teixeira et al., 2008); here brand presence increases attentional dispersion and promotes zapping.

There is relatively little research into the effects of other advertising elements than the ones described above. Pieters et al. (2007) show positive effects of *price and promotion elements* on attention to feature ads. In feature advertising these elements are crucial, and in optimal ads they should be larger than they currently are (about 60% for price, 10% for promotion). Most of the results of surface size effects on attention obtained for print and feature ads may not be generalized to other marketing stimuli, such as catalog pages, shelves, coupons, feature ads, billboards, and banner ads, although the study by Dreze and Hussherr (2003) affords useful initial insights in the latter. In addition, a recent study by Chandon et al. (2008) shows that increasing the number of facings on a brand on a shelf may promote attention by as much as 25%.

The editorial context and *competitive configuration* of these different marketing stimuli may also be very different, and as yet little is known on its effects. The work by Janiszewski (1998) first revealed that competition for attention plays a major role and may moderate the effects of surface sizes. Fox et al. (1998) showed that the ad context strongly affects attention to health warnings, and Wedel and Pieters (2000) establish the effects of context (media planning) variables, page location, and ad serial position. Consistent with this, Pieters and Wedel (2004) found negative effects of page location on attention to the ad as a whole, which may point to the negative effects of competition or fatigue.

As mentioned above, feature ads provide a unique opportunity to study visual competition, with multiple ads appearing on a single (newspaper) page. Based on Attention Engagement Theory (Duncan and Humphreys, 1992), Pieters et al. (2007) propose and investigate two measures of competitive clutter measures: target distinctiveness (TD) and distractor heterogeneity (DH). The former measures (in terms of entropy of element-surface size distributions) how different a feature ad is from all other ads on a page, the latter measures how different the other feature ads are from each other. Feature ads with a more distinctive design (higher TD) are selected more often. The more heterogeneous the other ads are (higher DH) the less frequently a target ad is selected. The duration of attention is influenced only by the distinctiveness of the ad-layout, distractor heterogeneity no longer plays a role. These effects of competitive clutter on attention present a critical trade-off in designing feature ad-displays: increasing the distinctiveness of a particular feature ad should improve both its attention selection and engagement, but simultaneously doing so for all feature ads in an ad-display increases distractor heterogeneity and thus hampers visual selection of the ads. This trade-off necessitates the use of optimization techniques described before.

These studies provide a starting point for more comprehensive investigation into competitive context and its effects on attention in a wider range of marketing settings. The study of billboards (d'Ydewalle and Tamsin, 1993) is of interest, because the competitive context is nearly absent, and it may provide insights into the effects of flow-movement (where the human body moves relative to the stimulus), and into vergence eye movements, in which both eyes move inwards to focus on an object that comes nearer. Furthermore, recently the study of the effect of competitive context of supermarket shelves has received interest (Van der Lans et al., 2008b; Chandon et al., 2008).

7.1.2 Low-level Perceptual Features

Relatively fewer studies have investigated the effect of other low-level features in visual exploration tasks. The studies by Lohse (1997), Lohse and Wu (2003), Rosbergen et al. (1997), Goldberg (1999), Van der Lans et al. (2008a,b) and d'Ydewalle et al. (1998) are exceptions. Rather than systematic inquiries, many of these studies include a few selected features, and mostly lack strong theoretical explanation of the effects found. These studies do demonstrate the effect of *color and edges* (line thickness) on visual attention. In the study by Rosbergen and coauthors this only occurred for one segment, i.e., 25% of the sample, and attention was higher for black and white than for color. Lohse (1997) however, finds that respondents view yellow-page color over 20% longer than ads without color and that finding was confirmed in Lohse and Wu (2001) Lohse and Wu (2001) for Chinese respondents, be it that the effect is somewhat smaller (11%). The study on search for health labels by Laughery et al. (1993) also assessed the effects of color, but did not find effects, nor did the study by Zhang et al. (2008) find that effect for feature ads. Goldberg et al. (1999) shows that thin lines help the findability of information on nutrition labels. The results together provide support for the effects of low-level features on visual attention to print ads, but because the effects are not very consistent across studies more research is needed. The findings of these various studies are not very easy to reconcile, but suggest that perhaps color contrast is more important than color *per se*. None of these studies addressed the basic low-level features that have been identified in basic attention research (Wolfe and Horowitz, 2004), including color, luminance, edges simultaneously, or explicitly studied the resulting salience.

The most extensive studies on low-level feature effects are those by Van der Lans et al. (2008a,b) for visual search of brands on shelves. They demonstrate the effects of color and edges in shelf displays on attention and search performance. Through the estimation of salience maps they demonstrate that colors and edges have strong effects on both localization and identification of brands on shelves. Which colors are salient depends in part on the search goal and varies substantially across individuals, but colors that are diagnostic in a search task (i.e., the color blue in a package) attract most attention. The results indicate, however, that consumers use only one or two basic features at the same time when trying to find a brand rapidly and accurately. The studies also suggest that individuals actively direct their focus of attention away from the edges and to the center of objects (brand groups), where diagnostic information is expected to reside. In the studies, these bottom-up effects explain about 2/3 of salience (top-down effects explaining the other 1/3). The authors use the salience maps to infer how brands and SKUs compete on shelves in terms of their visual salience, and demonstrate that salience is predictive of search performance. This competitive salience analysis reveals how brands can gain attention on the shelf by selectively changing their packaging designs.

7.1.3 Position

Position effects have been demonstrated in several studies, for example by Leven (1991), Rosbergen et al. (1997) for elements of ads, Goldberg (1999) for nutrition labels, Dreze and Hussherr (2003) for the placement of banner ads, d'Ydewalle and Tamsin (1993) for the placement of billboards during soccer games and Van Der Lans et al. (2008) and Chandon et al. (2008) for placement of brands on shelves. These studies found strong effects of (top/center/left) spatial location on attention. Goldberg et al. (1999) find that information in the center of a nutrition label is more difficult to find, because information is more dense, as compared to the top or bottom of the label. d'Ydewalle and Tamsin (1993) find positive effects for a central position of billboards on the TV screen, but only if (soccer) action is otherwise absent. Van der Lans et al. (2008a) show for search on shelves that the focus of attention is guided by systematic search strategies (left-right and right-left), which suggests that attention is guidance by the scene's organization. Chandon et al. (2008) find positive effects of central and top shelf placement on attention. Liechty et al. (2003) investigate the spatial distribution of attention to print ads by statistically distinguishing between a local (fixations on neighboring cells of a spatial grid) and global (fixations on non-neighboring cells) attention state, and show that subjects generally start in the local state and switch about 5 times between the two states during exposure. They found the intensity of the global state to increase over time, and found subjects to mostly end in the global state. The authors concluded that the scan path serves to break down the complex task of ad perception into a smaller number of simpler tasks of perception of local regions in the ad, where in the global state attention is redirected through peripheral vision between regions that are subsequently locally explored through foveal vision. Brasel and Gips (2007) find that during fast-forwarding brands located in the center of the screen receive more attention, but Teixeira et al. (2008) find that when brands are located in the center of the screen during regular viewing, TV commercials are zapped more often. Dreze and Hussherr (2003) find positive effects of a central location of banner ads on web-pages.

7.1.4 Movement

The study by d'Ydewalle et al. (1998) is important as one of the few studies addressing aspects of *movement* in dynamic stimuli. Their finding that shot-transitions affect perceptions of movement are of great interest. Clearly, the study of the effects of movement — as one of the basic perceptual features — on visual attention is in need of more inquiry, not only for TV commercials, but for example also for animated banner ads and clips. Teixeira et al. (2008) investigate the dynamic effects of cuts and edits as specific instances of pacing of TV commercials on attention and zapping, but do not find strong effects. Brasel and Gips (2007) show that viewers that fast forward pay more attention during commercials, but their attention is focused on the center of the screen. Fast-forwarded ads containing brand information in the center are therefore recognized after exposure, but ads with the brand located elsewhere are not.

7.1.5 Global Image and Ad Characteristics

Research has begun to examine the effects of characteristics of ad-images (or web Images) as a whole on attention to them. In a first study, Pieters et al. (2008) find that objective *visual complexity* increases attention to the ad. They also show that informativeness of the message increases attention to the ad. The attention lift was monotonic across the observed complexity range without evidence for diminishing returns and was mostly due to increased attention to the brand and pictorial elements. Attention to the ad increased with about 25%, from the minimum to the maximum complexity of ads in the sample. In their study, *incongruity* (artful deviations from expectations about the content of ads, such as impossible sizes, locations or features of objects) did not affect attention to ads. This finding counters common views in advertising that incongruity — a common tool to create original ads — increases the stopping power of advertising. Replication of this finding in future research is therefore called for. Pieters et al. (2002) studied the effect of *ad originality* as a more general characteristic of the ad as a whole. They found that original ads receive more attention (to the brand element) and are remembered better, even for the garden-variety originality defined in their study through a median split of the sample of ads. Pieters et al. (2008) find that message *informativeness* increased the attention of the ad as well, with over 25% from the minimum to maximum informativeness of the ads in the sample, mostly caused by increased attention to the pictorial and brand in the ad. Thus, adding "news" and other information cues in ads appears a viable strategy. For TV commercials, Teixeira et al. (2008) find that both under low and high levels of visual complexity the likelihood to stop watching commercials is higher than at intermediate levels of visual complexity, the density of visual information in the GIF-compressed file size of each frame in the commercials.

7.1.6 Optimizing Visual Stimuli for Maximal Effect

The efficiency and effectiveness of marketing expenditures can be improved by optimization of ad design, in terms of their features, such as the size of the ad elements. Such optimizations are challenging because they need to be based on an appropriate statistical model of gaze. Many of the attention models used are (log-) linear (e.g., Pieters and Wedel, 2004) and these models lead to so called "corner solutions," in which the element with the largest surface size effect will be the only element in the optimal ad. A recent study that has addressed that problem is Pieters et al. (2007). These authors use a hierarchical multivariate (Bayesian) regression model to explain attention to feature ads. As mentioned above, in their model they account for visual competition of feature ads on the display page by including measures of target distinctiveness and distractor heterogeneity, based on the theory of Duncan and Humphreys (1992). Their accounting for competition ensured internal solutions to the ad optimization problem. Pieters et al. optimize the layout of around 100 ad-display pages with multiple feature ads, taking into account the uncertainty in the model estimates. The feature ads are optimized in terms of the distribution of the sizes of their elements (brand, text, pictorial, price, and promotion), to maximize gaze on the entire ad display page. The results show that substantial gains in attention can be achieved, by reducing the size of the pictorial and text, and increasing the size of the brand and price, and to a lesser extent promotion elements, relative to current practice. The authors show that both manufacturers and retailers benefit from the optimization.

Teixeira et al. (2008) optimize the frequency and duration of placements of the brand in TV-commercials to minimize zapping, controlling for attention processes. Their results reveal that a pulsing strategy, with frequent but brief brand placements is optimal for many commercials. It is desirable to study the optimal design process for other visual marketing stimuli as well, including packages, shelves, print ads, and websites, since it has crucial importance for the practice of visual marketing.

7.1.7 Future Research

It would be of great interest to extend the studies on bottom-up effects summarized above further as many issues remain unexplored. Catalogs, feature ads, shelves, and web-pages, given their high level of concurrent competitive clutter, are appropriate venues to investigate the effects of clutter and context. Next to that, the further study on structural objects (text, pictorial and brand) in other the visual marketing stimuli than print ads would enable generalization of the effects found for print and feature advertising. The study of natural objects in these scenes, such as people and their faces, and of global image/scene characteristics is urgently needed. With the exception of the few studies summarized above, very little is known on these possible determinants of attention.

Because marketing stimuli often come in an editorial environment, e.g., in newspapers, magazines, television programs, and on web-pages, future research on attention competition and cooperation between marketing stimuli and their environment is needed. This is important in particular because media costs usually depend on the editorial environment in which ads and brands are placed, and remarkably little hard evidence about the positive or negative influence of editorial environments on attention to the ads and brands is currently available, thus leaving marketing practice to base their decisions on speculation and intuition. Exceptions are the studies by Fletcher et al. (1995) and Fox et al. (1999), who found strong effects of the advertising context on health warnings.

More insight is also needed about attention to dynamic stimuli, including TV commercials (but see Brasel and Gips (2007), and Teixeira et al. (2008)). Dynamic stimuli pose serious challenges to the conceptual and analytic models of attention, because the eyes moves in time and space, and the stimuli do so two, making eye-tracking data of dynamic stimuli "doubly-dynamic." Also given the rapidly growing use of dynamic stimuli on the web (clips, commercials, animated banners) this appears to be a fruitful area of future research. Finally, more research is needed to address the effects of the informational and emotional content of visual marketing stimuli on attention and downstream effects.

7.2 Person (Top-Down) Effects on Attention

7.2.1 Demographics

Demographic effects on visual attention have not yet been extensively studied. In the studies that have investigated *gender*, no or only small differences were observed (Dreze and Hussherr, 2003; Fox et al., 1998). Dreze and Hussherr (2003) are the only researchers that have studied *age*, and found an increase in gaze duration with age, presumably caused by slower information processing. In the only relevant *cross-cultural* research that we are aware of, no differences in attention patterns between cultures — United States and China — were found (Lohse, 1997; Lohse and Wu, 2003). Obviously, these initial studies call for follow-up research in which theories of age-related processing and memory changes, and cross-cultural differences in processing are tested.

7.2.2 Familiarity

Probably the most intensely studied top-down effect (in the context of print ads) on visual attention is brand familiarity. Russo and Leclerc (1994) showed that greater familiarity leads to more efficient choice of brands from shelves. A number of studies have revealed familiarity effects on attention to the ad as a whole, with differential effects of attention to ad objects (Pieters et al., 1999; Pieters and Wedel, 2004;

Table 7.2 Summary of top-down effects.

	Top-down effects				
Authors	Demographics	Familiarity	Involvement	Goals	Heterogeneity
Aribarg et al. (2008)		+	+		+
Dreze and Hussherr (2003)	+		+		
Fox et al. (1998)	+				
Goldberg et al. (1998)		+			
Janiszewski and Warlop (1993)		+			
Van der Lans et al. (2008a)					+
Van der Lans et al. (2008b)				+	+
Lohse (1997)	+				
Lohse and Wu (2003)	+				
Pieters et al. (2002)		+			+
Pieters et al. (1999)		+			
Pieters and Warlop (1999)			+		
Pieters and Wedel (2004)		+	+		
Pieters and Wedel (2007a)				+	
Rosbergen et al. (1997)		+	+		+
Russo and Leclerc (1994)		+			
Treistman and Gregg (1979)		+	+		
Rayner et al. (2001)				+	
Rayner et al. (2008)				+	
Wedel and Pieters (2000)					+
Wedel et al. (2008)				+	

Rosbergen et al., 1998; Treistman and Gregg, 1979). Familiar ads require and attract less attention than unfamiliar ads because they are easier to process. Pieters et al. (1999) found that across three exposures to print ads, the overall amount of attention dropped more than 50%, and other studies that use familiarity ratings confirm this for a variety of different attention measures. Familiarity decreases attention to the brand, does not affect pictorial attention, but increases attention to the text (Pieters and Wedel, 2004). The negative effect of familiarity can be mitigated by originality: for original ads the attention decrement is less as ads become more familiar (Pieters et al., 2002). The effect of familiarity is related to the effects of associative memory on attention, which was convincingly demonstrated by Janiszewski and Warlop (1993). The evidence provided by their studies suggest that brand conditioning creates a learning environment that makes semantic information available from memory and at the same time promotes attention to a brand. This may, in part, explain the familiarity effect.

7.2.3 Expertise, Involvement, and Attitudes

Two studies investigating the effect of *expertize* or training in quite different contexts, report comparable effects (Dreze and Hussherr, 2003; Goldberg et al., 1999). With increasing experience gaze durations in self-paced exposure conditions decreases, but the scan paths seem to become more regular. Although mostly used as a control variable, a number of studies find positive effects of *involvement* on attention (Rosbergen et al., 1997; Treistman and Gregg, 1979) but in the large scale study by Pieters and Wedel (2004) the effect was not found. The aggregation of the data across participants, or the measurement of involvement by external judges — rather than by the participants — at the level of ads, may have masked the effect in that study. Pieters and Warlop (1999) did find more extensive information processing under high *motivation* (involvement) conditions, but less intensive processing of pictorial information. Interestingly they also find and increase in the average duration of fixations, which is consistent with research by Goldberg et al. (1999) and Rayner et al. (2001) on differences between information types (text versus pictorials). One possible explanation of the time effect may come from the research by Liechty et al. (2003). High motivation conditions may affect the time individuals spend in the "local" covert attention state. This is supported by the result of Pieters and Warlop that participants made less between-brand saccades, which would point to a reduced amount of time in the "global" state. Aribarg et al. (2008) confirmed that more positive *brand attitude* increases attention to the brand and the text, but did not confirm the effect on fixation durations. The measurement instrument used to assess attitudes, and the study context and stimuli used may play a role, but clearly more research is needed to enable generalization of these findings. In their study, consumers also claimed recognition memory for ads of well-liked brands over and above the actual attention they spend to the ads during exposure, which revealed a bias in recognition measures. Pieters and Warlop (1999) also found effects of *time pressure*, which directly affects the extent of information processing and has more or less opposite effects of those of motivation. Under pressure, in particular textual information was skipped, which may be explained from the fact that text requires more intense processing.

7.2.4 Goals

Research on the influence of consumer *goals* on the processing of visual marketing stimuli such as print ads and product displays has only recently begun. Studies by Pieters and Wedel (2007a) and Rayner et al. (2008), have shown the influence of scene perception goals on spatial patterns of eye movements on ads, that is, where people look and for how long. Wedel et al. (2008) study the effects of goals on the time course of visuo-spatial attention, Chandon et al. (2008) investigate the effect of shopping goals on attention to shelves. Rayner et al. (2001) and Rayner et al. (2008) found different patterns of attention to ads, depending on whether respondents were asked to imagine wanting to buy the brand advertised or asked to rate the ads for likability or effectiveness. In the latter study/condition participants looked more at the pictorial in the former more at the text of the ad. These findings are also discussed in Rayner and Castellano (2007).

Building on the work by Yarbus (1967), Pieters and Wedel (2007a,b) found different patterns of attention for each of four processing goals, as compared to a fifth condition of free viewing. They found that the gaze on the elements of ads, adjusted for their sizes, was similar under an ad-appreciation goal. Under a brand-evaluation goal, people dwelled on the text longer, but under a brand-learning goal they dwelled more on the text, but less on the pictorial. Text, pictorial, and brand elements were all looked at longer under an ad-memorization goal. Differences between goals were largest for the body text and smallest for the headline which was equally informative under all goals. The authors did not observe goal effects on pupil diameter. The attention patterns for free viewing and the ad-appreciation goal were brief and remarkably similar, which may point to ad-appreciation being the dominant goal under free-viewing. Figure 7.2 summarizes these effects of goals on the duration of attention to structural objects in advertisements. The different gray-scale bars distinguish attention due to each of the four

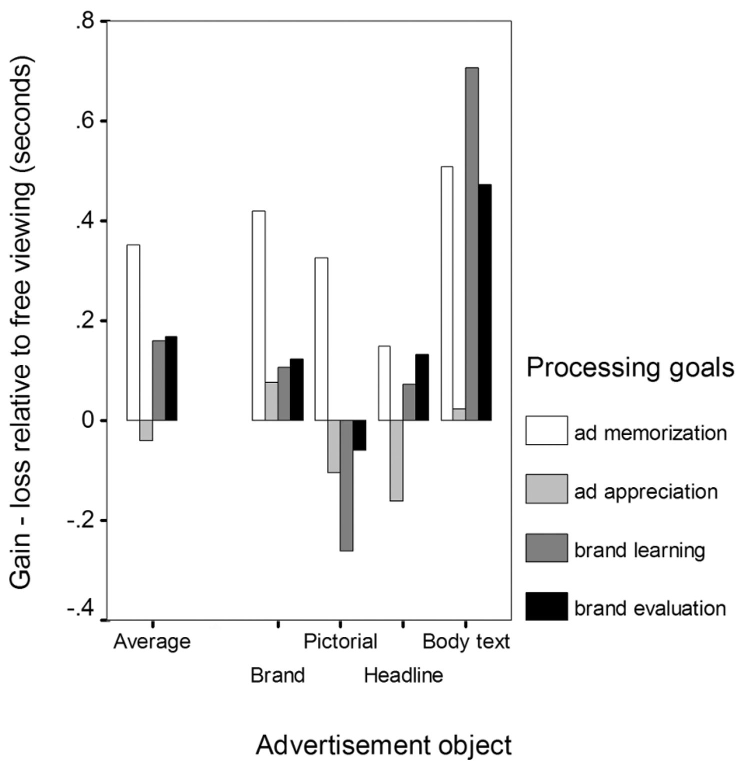

Fig. 7.2 Goal control of attention duration: Gains and losses relative to the free viewing condition: (Reproduced from Pieters and Wedel (2007a,b)).

examined goals, relative to the free viewing condition. Different groups of bars reflect each of the four structural ad objects, brand, pictorial, headline, body text that were examined. The large differences between goals and the relative suppression of the pictorial under all goals except ad memorization are striking. The findings of this study reflect the specificity of the effects of goals on attention to ads and information acquisition strategies (see also Rayner et al. (2008)). They are important, because the same set of advertisements was explored under different goals manipulated between-subjects. Thus, differences in attention patterns to ads and their objects can be unequivocally attributed to differences in the informativeness of the ad objects, rather than to differences in their salience, because the latter was held constant.

Wedel et al. (2008) show that these scene perception goals also systematically influence the frequency of switching between the local and global attention state described previously (see Liechty et al. (2003)). Goals directed at processing the ad as a whole led to higher frequencies of switching between states than did goals directed at processing a specific target (the brand) in the ad, which resulted in the longest duration of the local state. Goals directed at the ad as a whole thus require more complete examination of the ad, which is handled by breaking up the scene in spatially dispersed areas of interest for local exploration. Evaluation goals led to shorter durations of the local state than each of their learning goal counterparts. This shows that evaluation goals can generally be attained faster. This study complements the earlier study by Pieters and Wedel (2007a) by providing a spatial rather than an object-based account of the goal control of visual attention, using a qualitatively different analytic approach to closely analyze some of the same data. Figure 7.3 summarizes the key findings. As the figure demonstrates, each of the four goals and free viewing led to a remarkably distinct patterns of duration in the local attention (horizontal dimension), which is assumed to be involved in detailed information intake, and to different switching frequencies between this state and the global attention state (vertical dimension). The findings are noteworthy because they demonstrate how goals led to qualitatively different attention patterns, which persisted across a larger set of 17 advertisements in their regular magazine context.

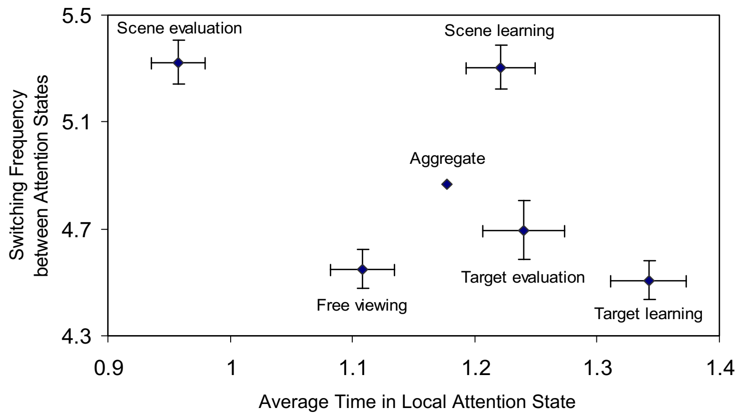

Fig. 7.3 Goal control of latent attention switching: (Reproduced from Wedel et al. (2008)).

Jointly, the findings reveal the systematic effects of goals on overt and covert attention patterns, and how the informativeness of advertisements and their objects is not an intrinsic characteristic of the ads, but the multiplicative effect of top-down and bottom-up characteristics. This means that research on the information value of advertising needs to take the goals of consumers into account, to show not *if* but *when* ads are informative and for *which* goals.

Van der Lans et al. (2008b) investigated the effects of search goals on attention during target search tasks for brands on shelves. Target search is one of the most extensively studied tasks in psychology and engineering, and is one of the most frequent tasks that consumers engage in, when searching for products and brands in retail stores, catalogs, and websites. The authors estimated salience maps of search displays under different search tasks. Their results reveal dramatic effects of search goal effects on brand salience, in terms of the basic features that are enhanced or inhibited during search. They found that about 1/3 of salience on the shelf is due to these top-down effects, which component of brand salience can be influenced through out-of-store advertising. The authors derived the implications of their findings for the optimum visual differentiation level of brands and their SKUs.

We see the findings on the effects of goals on attention to print ads and shelves during search as a first step. In these first few studies, goals have been shown to have pervasive effects on attention, but only a few goals (processing and search goals) have been investigated, and the effects of many other goals remain to be charted. Further, goal effects on attention to other marketing stimuli, in particular web-sites, where browsing and buying goals have been demonstrated to affect the click-stream is welcome.

7.2.5 Individual Differences

Some studies have taken unobserved differences between consumers into account in statistical models of attention (Aribarg et al., 2008; Pieters et al., 2001; Rosbergen et al., 1997; Van der Lans et al., 2008a,b; Wedel and Pieters, 2000). These unobserved differences between consumers may be due to heterogeneity in the influence of top-down factors operating between them. A potentially interesting finding from two of these studies (Wedel and Pieters, 2000; Pieters et al., 2001) is that heterogeneity between ads seems to be greater than heterogeneity between consumers, potentially indicating that under natural exposure conditions to print ads, bottom-up influences may be larger than top-down influences. The studies by Van der Lans et al. (2008a,b) and Chandon et al. (2008) confirm this for attention during shelf search, where two-thirds of the variability was accounted for by bottom-up factors. This would also accord with the stability of scan-paths that has been found (Harris, 1993; Pieters et al., 1999), presumably also caused by dominant effects of bottom-up factors residing in the ads. Finally, Rosbergen et al. (1997) find three unobserved groups of people that have a different distribution of attention across the elements of print ads. These segments may arise due to different effects of top-down factors, possibly goals that participants have while exploring the ads.

7.2.6 Cognition

Previous eye-tracking studies have often been descriptive, which is already interesting, but leaves the potential of eye-movement data as information processing data untapped (Russo, 1978). The more recent research in the last decade attempts to infer covert cognitive processes from eye-tracking data. Pieters et al. (1999) test the postulates of Noton and Stark (1971) and show the stability of scan-paths of eye movements on images, across consumers, and repeated exposures. Josephson and Holmes (2002) replicate that study for web-pages and find similar results.

Both in exploration and search tasks, researchers have found cognitive states to drive the overt visual exploration of the scene. Russo and Leclerc (1994) find three different states, orientation, evaluation, and verification. But these authors did not statistically test for these states. Liechty et al. (2003) and Wedel et al. (2008) postulate two states, local and global attention, and show that this provides a much better representation of the eye-movement data than a single state, but do not test for three or more states.

Strictly speaking, the results of these studies are difficult to compare, since Russo and Leclerc use a shelf search/choice task, and Liechty et al. and Wedel et al. used visual exploration and goal directed tasks for print ads. Nevertheless, there is some correspondence between the states, the Russo and Leclerc evaluation and verification states show similarity with Liechty and coauthors' local attention state, and the exploration and global attention states also have communalities. Van der Lans et al. (2008a,b) do investigate shelf search and show that similar unobserved attention states, which they call localization and identification, play a role during search. The localization state is driven by a salience map of the visual field and systematic search patterns, the identification state is driven by re-fixations on perceptual groups and objects.

An important insight of all these studies afford is that rather than constituting a strict linear and unidirectional sequence from exploration or localization (global) to verification or identification (local), the activity of mental states rapidly and repeatedly switches back and forth over time, even during very brief visual exposures. Although the postulate of discrete mental states underlying cognition and behavior has played a key role in psychological theory and has been vital in the study of temporal properties of mental processes, there has been an intense debate on whether such states exist, or that cognition is better represented as a continuous process (see, e.g., Yantis et al., 1991). This parallels the discussion on discrete versus continuous heterogeneity in marketing (see, e.g., Wedel et al., 1999). The final verdict on the number of states is still out, and the extent to which states adequately represent the cognitive process and the role of the interplay of bottom-up and top-down factors on the activation of these states needs to be further investigated.

7.3 Information in Pupillary Diameter

Pupil diameter has been associated with specific emotional (more general: evaluative) responses and with cognitive responses, mostly of cognitive load or intensity of processing, and it is relevant to distinguish these two and examine them in more detail. Although research on emotional responses and pupillary response dates back several centuries (Loewenfeld, 1999), the field was re-opened by Hess and Polt (1960), and much of the recent ideas and work builds on their claims. They explained that "a man's pupils will enlarge promptly and substantially when he is looking at a comely girl. On the other hand, the pupils will contract when something presumably distasteful is in view." Thus, the claim was of a "bidirectional" pupillary response, such that pleasant ads led to pupil dilation and unpleasant ads to pupillary constriction. This idea and the data to support it led to an avalanche of publicity, excitement in the marketing and advertising community, that saw an objective (nonverbal) and specific (bi-directional) indicator on emotional response at hand (Rice, 1974). Krugman (1964) published a supportive article in the first volume of the *Journal of Marketing Research*. Loewenfeld (1999) in her definitive work on the pupil reports on over 100 studies that were conducted between 1960 and 1982 on the issue. She concludes (p. 667): "Now, about 25 years since Hess' first publications, what has been accomplished by all this expenditure of work and time? Nothing, really. It has been shown over and over again that what could not be, according to the anatomic and physiologic properties of the iris system, really was not: emotional stimuli and all other sensory and psychologic stimuli — with the exception of light, and stimuli that alter the eye's near point of vision — do not constrict the eye but dilate it." Thus, it is not the case that emotional stimuli do *not* affect pupil diameter. But and importantly, there is no bi-directional relationship, and emotional stimuli of positive and negative valence both dilate the pupil. This significantly reduces the value of pupil dilation for emotion measurement, since the same pupil dilation can mean joy or disgust, and academics and advertisers may like to distinguish these two diverse responses. Moreover, the strong emotional processes that may prompt pupil dilation are most likely accompanied by increased arousal and processing load, and these may actually be the responsible processes, which further lowers the validity of pupil diameter as a measure of emotional response. Despite the limited empirical support for a clear bidirectional relationship between pupil diameter and emotional responses, and the general paucity of research in the last decades, websites of ad agencies, marketing research books, methodology chapters in general marketing books, and even psychophysiology books (e.g., Andreasi, 2000) continue to uncritically laud the utility of pupil dilation as a specific measure of emotional processing, and of ad and brand preference.

There is more support for a systematic relationship between increased cognitive load and pupil dilation, which is termed the "cognitive activation" hypothesis (Beatty and Lucero-Wagoner, 2000). Loewenfeld (1999) reports almost 100 studies conducted between 1964 and 1982 relating aspects such as increasing task demands, intensity of information processing, effort, cognitive load, and their predictive value in memory and other tasks, that support the hypothesis in varying degrees. However, care in interpretation is needed here as well. First, if increasing task demands lead to increased pupil dilation, the question is whether the pupil dilation is due to the increased task demands or to the increased cognitive resources that are allocated to deal with the task demands, and the interpretation of results varies clearly depending on the inference (Matthews et al., 1991). Thus, Pieters and Wedel (2007a) found significant differences between processing goals (with different attentional demands) on gaze durations to specific objects in print ads, but no differences in pupil diameter, perhaps because people looked longer rather than more intensely. Second, pupil dilation has also been found in auditory tasks (without a visual stimulus), and in mental arithmetic and imagery tasks that vary in demands (without an external stimulus altogether) (Kahneman, 1973). The question then becomes what the causal links between pupil diameter and cognitive processes are. In a recent review of the literature Beatty and Lucero-Wagoner (2000, p. 143) conclude that pupillary responses are best seen as "reporter variables" that "are unrelated in any obvious way to central cognitive processes, yet [they] nonetheless empirically reflect variations in central processing load with extraordinary precision." A recent example is a study reporting reliable effects on pupil dilation of decisions to deceive other players in an experimental economics game (Tao-yi Wang et al., 2008). Although the findings could be interpreted as effects of emotional responses — due to lying, they are most likely the result of the increased cognitive load — due to the unnatural decision to deceive.

Clearly, further marketing research under controlled conditions and with appropriate statistical controls, as we advocate here, is needed to make valid inferences about the influence of task characteristics such as task complexity, and person characteristics such as amount of cognitive resources allocated. Yet, in view of the complexities of the sources of intra-ocular eye movements, we believe that a focus on extra-ocular eye movements should lead to more progress in the short-run. The models and methods that have been developed to understand extra-ocular eye movements (fixations and saccades) may prove useful in combination with measures of Task-Evoked Pupillary Responses (TEPR) to understand intra-ocular eye movements may lead to new findings and theories, that are of relevance to visual marketing.

7.4 Downstream Effects

There is substantial evidence that even small attentional effects on print ads and other commercial visual stimuli, such as catalogues, shelves, labels, and packaging designs, affect memory (Fisher et al., 1989; Fletcher et al., 1995; Janiszewski, 1998; Krugman et al., 1994; Pieters and Warlop, 1999; Pieters et al., 2001; Rosbergen et al., 1997; Treistman and Gregg, 1979; Wedel and Pieters, 2000; Witt, 1977). This persistent association supports the validity of eye-tracking measures of visual attention. Aribarg et al. (2008) show, however, that memory-based ad-recognition measures, such as the Starch-measures often used in advertising practice, are associated with attention to the ad, but are subject to biases. Although the authors show how these biases can be corrected, their study leads to the conclusion that it is unwise to use these measures to substitute for actual attention.

Further downstream effects of attention have been identified as well, such as positive effects on attitudes, preferences, and intentions (Pieters and Warlop, 1999; Rosbergen et al., 1998; Treistman and Gregg, 1979). Chandon et al. (2007) and Chandon et al. (2008) show effects of visual lift on consideration-set formation in a point-of-purchase setting. Van der Lans et al. (2008a,b) show strong effects of attention on search performance — latency and accuracy — on shelves. An indication that attention to print ads may lead to more sales was provided by Treistman and Gregg (1979), who observed that one of two ads that people looked at longer in an eye-tracking experiment also received more sales. Janiszewski (1998) was not able to establish an effect of attention on sales, over and above what was explained by the surface size effect. Zhang et al. (2008) show for feature advertising that attention to the ads predicts subsequent retail sales, even while controlling for other potential determinants. Their study reveals that gaze duration mediates the effect of feature ad characteristics (sizes, color, and location) on sales. In establishing that mediating relationship, they find and account for endogeneity of surface sizes. Endogeneity of surface size in their attention model suggests that manufactures and/or retailers may strategically determine surface sizes based on a (proxi-) measure of attention. If that finding holds for other visual marketing stimuli as well, in particular for print ads, the effects of surface sizes may be even larger than believed to date. Although the study by Zhang et al. is important because it is the first to establish the relation between attention and sales, the feature advertising context is specific and most conducive to uncovering such a relationship because of the immediate effects of feature ads. There is a great need both for academic and practical purposes to more firmly establish the attention-sales relationship across a wider range of visual marketing stimuli and consumer decision contexts.

8

Conclusion

8.1 Recommendations for Marketing Practice

Which lessons can be learned from eye-tracking research in marketing? Below are a number of stylized recommendations derived from research findings to date. The amount of research on which these findings are based varies, so that some recommendations can be stronger than others.

Differences in attention patterns between visual stimuli are generally to be greater than between consumers, potentially indicating that under natural exposure conditions to print ads, ad-specific or "bottom-up" influences may be larger than consumer specific or "top-down influences," a rough estimate being that they account for about 2/3 of the variability in attention patterns.

8.1.1 Print Advertising: Size Matters in Capturing Attention

There are four common design elements for any print ad: brand, text, pictorial, and headline. Much is known on the effects of surface size of these elements of print ads, which leads to the following.

(1) *Brand/logo.* Attention to the brand increases modestly with 0.3% with a 1% increase in surface size. The brand element should be played up because it plays crucial role in routing attention through ads and improves memory for the advertised brand.
(2) *Text.* Increase in attention to the text in ads is almost proportional: a 1% increase in surface size of the text leads to a 0.9% increase in gaze, but there is doubt on the central role of text in print ads for integrating the message and subsequent brand memory.
(3) *Pictorial.* The pictorial attracts attention almost regardless of its size and advertisers would be ill advised making the pictorial large regardless of its content.
(4) *Headline.* Receives consistent attention independent of consumer goals. Top-location and large font size improve attention.
(5) *Color and Edges.* Color may have positive effects on attention, but depending on context attention may be higher for black-and-white than for color. Most likely color contrast matters most.

8.1.2 Feature Advertising: Optimal Organization of Ads Improves Attention

There are five common design elements for any feature ad (FSI, Free Standing Insert), the first three are the same as for print ads: brand, text, and pictorial. But, in addition, feature ads often have distinct price and promotion elements. Feature advertising is particularly susceptible to the adverse effects of competitive clutter, much more so than other forms of print advertising, because multiple feature ads appear on the same ad-display page. The following recommendations can be made.

(1) *Element Sizes.* Pictorial and text elements often could be smaller by 35 an 20% respectively, while the brand, price, and promotion elements could be increased in size by 75%, 60%, and 10% respectively, relative to current practice. The optimal layout for manufacturer and private label ads is

somewhat different, where the former should have somewhat larger pictorial and promotion elements, but less text (5%–10%).

(2) *Competitive Clutter*. A uniform layout of feature ads on a page reduces clutter and improves attention to each of the ads. More ads on a page reduces attention to each of them, and in addition makes their pictorial elements less effective. Feature ads with a more distinctive design are selected more often visually than those that are similar to other ads, but the more different the other ads are the less frequently a target ad is selected.

8.1.3 TV Commercials: Brand Pulsing Prevents Avoidance

Key aspects of the design of TV commercials is the presence of the brand. The following recommendations follow from research to date.

(1) *Concentrating visual attention* in TV commercials makes consumers stay with the commercial: the power to orchestrate attention is crucial for commercial effectiveness.
(2) A *brand pulsing* strategy where brands are shown more frequently but briefly instead of infrequently but longer, decreases commercial avoidance rates substantially.
(3) If exposure during *fast-forwarding* is required, brands should be located in the center of the screen. But a similar strategy may promote more zapping during regular exposure.

8.1.4 Banner Ads: Size and Centrality Promote Attention

Research is limited as yet, but the available findings suggest the following.

(1) *Larger* banners attract more attention and a more *central location* has a positive effect on attention.
(2) Banner ads may only be attended to in half of the cases or less, but may still have an effect through *incidental exposure*.

8.1.5 Labels and Warnings: Visual Layout Matters

The research on labels and warnings, although somewhat less recent, suggests the following.

(1) For nutrition labels, information near the (denser) *center of a label* may require more time to find than when at the top or bottom of the label. *Thinner alignment lines* result in faster search than thicker ones.
(2) *Pictorials, icons, and color* substantially improve noticability of health warnings, whereas borders around the warning may not have much effect.
(3) *Advertising and product context* has strong effects on labels health warnings.

8.1.6 Shelves: Placement and Optimal Visual Differentiation of SKUs and Brands is Critical

Surprisingly, relatively few studies have as yet addressed the effectiveness of in-store marketing efforts. The research is recent and leads to a number of interesting findings.

(1) Consumers use only *one or two basic features* at the same time when trying to find a brand and advertising needs to establish strong memory associations with a limited number of unique features.
(2) Visual *under-differentiation* of brands leads to brand confusion, *over-differentiating* the brand from the category may result in difficulties finding the brand.
(3) Visual *under-differentiation* of SKUs may lead to cannibalization, *over-differentiation* of SKUs may diminish unique brand associations and erode brand equity.
(4) A larger number of facings, and top and/or central placement increases attention to and consideration of the brand.
(5) If all brands on the shelf are impulse brand with few memory associations, the retailer should use a "stick with the winner" strategy. If all brands have strong memory associations, the

retailer should use a "help the poor" strategy of allocating in-store visual salience.

8.1.7 Top-Down (Person) Effects

Most of the research on top-down effects involves print ads, and much less is known about decision and choice effects. It leads to the following recommendations.

1. *Demographics.* Older people need more time to process ads, but cultural differences may be relatively small.
2. *Familiarity.* Familiar ads require and attract less attention than unfamiliar ads: attention may drop by as much as 50% at each subsequent exposure. As ads become familiar, attention to the brand decreases and attention to the text increases.
3. *Originality.* The negative effect of familiarity on attention can be mitigated by originality: for original ads the attention decrement is less as ads become more familiar, and memory is enhanced. Originality also leads to more attention in itself.
4. *Involvement and time pressure.* For high involvement categories and motivated consumers, pictorial information is less important, and for ads targeted to consumers or contexts with limited available time text is less important.
5. *Goals.* Stimulating a goal to remember the ad makes consumers pay more attention to the body text, pictorial, and brand. Stimulating brand memory makes consumers pay more attention to the text but less to the pictorial. Consumers pay more attention to the body text if they try to evaluate the brand. Priming goals through ads themselves is desirable and may have pervasive effects on attention patterns.

8.1.8 Downstream Effects: Attention Affects Memory, Preference, and Sales

Attention to print ads and other commercial visual stimuli, such as catalogues, shelves, labels and packaging designs, affect memory and preferences. Recognition memory such as in Starch scores, however, is not a good proxi for attention because it is subject to bias. Further downstream effects have been identified, including positive effects on attitudes, preferences, intentions, search performance, choice, and sales. Increasing attention to visual marketing stimuli is likely to improve sales and profitability.

8.2 Outlook

The last decade has seen strong integration of methods and theories from vision, attention, Bayesian statistics and eye-tracking research in order to inform and further improve marketing decision making. These developments are important, but earlier progressed rather disparately. Vision science offers the tools to address the extraction of basic perceptual features from images, image segmentation, object recognition, and so forth. Attention research offers theories that enable precise predictions of attention effects, and experiments that facilitate establishing causal relations between constructs. Eye tracking offers methods of recording, representing and interpreting eye movements. Bayesian statistics offers the tools to formalize these theories in realistic statistical models that enable inference on unobserved attention processes.

In future research on eye tracking, combining experimental procedures and statistical models will enable us to get a view of "the cognitive iceberg, from the visible tip of the eye movements" (Russo, 1978). Large samples of participants, but perhaps more importantly large samples of marketing stimuli will enable generalizations of findings that have an impact on marketing practice. It is a priority to develop integrative conceptual and statistical models of visual attention that include causal effects of perceptual and conceptual features, and represent the interplay of bottom-up and top-down effects on salience and informativeness of visual marketing stimuli. In order to gain increased understanding of visual attention in real-life, it may prove valuable to integrate statistical models and their estimates of key parameters of attentional processes, with cognitive simulation models such as EPIC (Kieras and Meyer, 1997), and its extensions. With key attention parameter estimates as inputs, such simulation models may allow for predictions from much richer models, firmly grounded in theories of attention and memory, that would be difficult to bring to empirical eye-movement data in all their detail. We see potential for the further development of formal methods to optimize the design of visual marketing stimuli, including print and feature ads, TV commercials, billboard, shelves, and packages. Finally, eye-movement patterns on a wider range of marketing stimuli are called for, with web-stimuli being a priority because of high external validity of eye-tracking experiments, dearth of current research, commercial importance, and multi-modal visual environments, including sponsored search, clips, banner ads, product reviews, product comparison lay-outs, tag-clouds, and editorial content.

It is evident that eye movements under normal conditions are strongly and directly connected to high-order cognitive processes (Findlay and Gilchrist, 1998, 2001; Fuster, 2003; Rizzolatti et al., 1994; Pieters and Wedel, 2007b). Eye movements may not have (yet) saved the world (Russo, 1978), but they can contribute significantly to marketing reaching its full potential, by providing an unprecedented view on consumers' moment-to-moment processing of visual marketing stimuli, with predictive validity for downstream effects. We hope that this survey will contribute to the wider use of eye-movements for inverse inferences on covert visual attention processes, and the further integration of vision science, attention theory, statistical modeling and eye-movement research. This will lead to a better understanding of visual attention to marketing stimuli in realistic settings, and thus to more effective visual marketing, because "without attention all else is hypothetical."

References

Andreasi, J. L. (2000), *Psychophysiology: Human Behavior & Physiological Response*. Mahwah, NJ: Lawrence Erlbaum.

Anstis, S. M. (1974), 'A chart demonstrating variations in acuity with retinal position'. *Vision Research* **14**, 589–592.

Antes, J. R. (1974), 'The time course of picture viewing'. *Journal of Experimental Psychology* **103**, 62–70.

Aribarg, A., R. Pieters, and M. Wedel (2008), *The Diagnostic Value of Advertising Recognition*. Working Paper, University of Michigan.

Beatty, J. and B. Lucero-Wagoner (2000), 'The pupillary system'. In: J. Cacioppo, L. G. Tassinary, and G. G. Berntson (eds.): *Handbook of Psychophysiology*. Cambridge: Cambridge University Press, pp. 142–162.

Berlyne, D. E. (1958), 'The influence of complexity and novelty in visual figures on orienting responses'. *Journal of Experimental Psychology* **55**(3), 289–296.

Berlyne, D. E. (1971), *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts.

Berlyne, D. E. (1974), *Studies in the New Experimental Aesthetics: Steps Toward an Objective Psychology of Aesthetic Appreciation*. New York: John Wiley & Sons.

Bettman, J. R., E. J. Johnson, and J. W. Payne (1991), 'Consumer decision making'. In: T. S. Roberson and H. H. Kassarjian (eds.): *Handbook of Consumer Behavior*. Englewood Cliffs, NJ: Prentice-Hall, pp. 50–84.

Brasel, S. A. and J. Gips (2007), 'Breaking through fast-forwarding: Brand information and visual attention'. Marketing Science Institute Special Report 07–201, September 2007.

Broadbent, D. E. (1958), *Perception and Communication*. London: Pergamon.

Burke, R. R. and T. K. Srull (1988), 'Competitive inference and consumer memory for advertising'. *Journal of Consumer Research* **15**(1), 55–68.

Buswell, G. T. (1935), *How People Look at Pictures: A Study of the Psychology of Perception in Art*. Chicago: University of Chicago Press.

Calvo, M. G. and P. J. Lang (2004), 'Gaze patterns when looking at emotional pictures: Motivationally biased attention'. *Motivation and Emotion* **28**(3), 221–243.

Chandon, P. (2002), 'Do we know what we look at? An eye-tracking study of visual attention and memory for brands at the point of purchase'. Working Paper, INSEAD, Fontainebleau.

Chandon, P., J. W. Hutchinson, E. T. Bradlow, and S. H. Young (2007), 'Measuring the value of point-of-purchase marketing with commercial eye-tracking data'. In: M. Wedel and R. Pieters (eds.): *Visual Marketing*. Lawrence Erlbaum, pp. 225–259.

Chandon, P., J. W. Hutchinson, E. T. Bradlow, and S. H. Young (2008), 'Does in-store marketing work? Effects of the number and position of shelf facings on attention and evaluation at the point of purchase'. Working Paper, The Wharton School, University of Pennsylvania.

Chen, Y. and S. Yang (2007), 'Estimating disaggregate models using aggregate data via augmentation of individual choice'. *Journal of Marketing Research* **4**, 596–613.

Corbetta, M. and G. L. Shulman (2002), 'Control of goal-directed and stimulus-driven attention in the brain'. *Nature Reviews Neuroscience* **3**, 201–215.

Deco and Zihl (2001), 'Top-down selective visual attention: A neurodynamical approach'. *Visual Cognition* **8**, 119–140.

Desimone, R. and J. Duncan (1995), 'Neural mechanisms of selective visual attention'. *Annual Review of Neuroscience* **18**, 193–222.

Deutsch, J. A. and D. Deutsch (1963), 'Attention: Some theoretical considerations'. *Psychological Review* **70**, 80–90.

Dodge, R. (1900), 'Visual perception during eye movements'. *Psychologcal Review* **7**, 454–465.

Donderi, D. (2006), 'Visual complexity: A review'. *Psychological Bulletin* **132**(1), 73–97.

Dreze, X. and F.-X. Hussherr (2003). 'Internet advertising: Is anybody watching?' *Journal of Interactive Marketing* **17**(4), 8–23.

Duchowski, A. T. (2002), 'A breadth-first survey of eye-tracking applications'. *Behavior Research Methods, Instruments, and Computers* **34**(4), 455–470.

Duchowski, A. T. (2003), *Eye Tracking Methodology: Theory and Practice*. New York: Springer–Verlag.

Duncan, J. (1980), 'The locus of interference in the perception of simultaneous stimuli'. *Psychological Review* **87**(3), 272–300.

Duncan, J. (1996), 'Cooperating brain systems in selective perception and action'. In: T. Inui and J. L. McLeland (eds.): *Attention and Performance XVI*. Cambridge: MIT Press, pp. 549–578.

Duncan, J. and G. W. Humphreys (1992), 'Beyond the search surface: Visual search and attentional engagement'. *Journal of Experimental Psychology: Human Perception and Performance* **18**(2), 578–588.

d'Ydewalle, G., G. Desmet, and J. Van Rensbergen (1998), 'Film perception: The processing of film cuts'. In: G. Underwood (ed.): *Eye Guidance in Reading and Scene Perception*. Amsterdam: Elsevier, pp. 357–368.

d'Ydewalle, G. and F. Tamsin (1993), 'On the visual processing and memory of incidental information: Advertising panels in soccer games'. *Visual Search* **2**, 401–408.

d'Ydewalle, G., P. VandenAbeele, J. Van Ronsgerger, and D. Concke (1988), 'Incidental processing of advertisements while watching soccer game broadcasts'. In: M. M. Gruneberg, P. E. Morris, and R. N. Sykes (eds.): *Practical Aspects of Memory: Current Research and Issues,* Vol. 1, Memory in Everyday Life. Chichester: Wiley, pp. 470–483.

Feng, G. (2004), 'From eye movement to cognition: Toward a general framework of inference'. *Psychometrika* **68**, 557–562. Comment on Liechty et al. 2003.

Findlay, J. M. and I. D. Gilchrist (1998), 'Eye guidance and visual search'. In: G. Underwood (ed.): *Eye Guidance in Reading and Scene Perception*. Amsterdam: Elsevier Science Ltd., pp. 295–312.

Findlay, J. M. and I. D. Gilchrist (2001), 'Visual attention: The active vision perspective'. In: M. Jenkin and L. Harris (eds.): *Vision and Attention*. New York: Springer Verlag, pp. 83–103.

Fischer, P. M., J. W. Richards, E. J. Berman, and D. M. Krugman (1989), 'Recall and eye tracking study of adolescents viewing tobacco advertisements'. *Journal of the American Medical Association* **261**(1), 84–89.

Fletcher, J. E., D. M. Krugman, R. J. Foit, P. Fischer, and T. Rojas (1995), 'Masked recall and eye tracking of adolescents exposed to cautionary notices in magazine ads'. In: P. S. Ellen and P. J. Kaufman (eds.): *Marketing and Public Policy Conference Proceedings*, Vol. 5. Atlanta: Georgia State University, pp. 128–135.

Folk, C. L., R. W. Remington, and J. C. Johnston (1992), 'Involuntary covert orienting is contingent on attentional control settings'. *Journal of Experimental Psychology: Human Perception and Performance* **18**, 1030–1044.

Fox, R. J., D. M. Krugman, J. E. Fletcher, and P. M. Fischer (1998), 'Adolescents' attention to beer and cigarette print ads and associated product warnings'. *Journal of Advertising* **27**(3), 57–68.

Friedman, A. (1979), 'Framing pictures: The role of knowledge in automatized encoding and memory for gist'. *Journal of Experimental Psychology: General* **108**(3), 316–355.

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

Geman, S. and D. Geman (1984), 'Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721–741.

Germeys, F. and G. d'Ydewalle (2007), 'The psychology of film: Perceiving beyond the cut'. *Psychological Research* **71**, 458–466.

Goldberg, J. H., C. K. Probart, and R. E. Zak (1999), 'Visual search of food nutrition labels'. *Human Factors* pp. 425–437.

Harris, C. M. (1993), 'On the reversibility of markov scanning in free viewing'. *Visual Search* **2**, 123–135.

Henderson, J. M. (2003), 'Human gaze control during real-world scene perception'. *Trends in Cognitive Sciences* **7**(11), 498–504.

Henderson, J. M. and A. Hollingworth (1998), 'Eye movements during scene viewing: An overview'. In: G. Underwood (ed.): *Eye Guidance in Reading and Scene Perception*. Amsterdam: Elsevier, pp. 269–293.

Hess, E. H. and J. M. Polt (1960), 'Pupil size as related to the interest value of visual stimuli'. *Science* **132**, 349–350.

Inhoff, A. W. and R. Radach (1998), 'Definition and computation of oculomotor measures in the study of cognitive processes'. In: G. Underwood (ed.): *Eye Guidance in Reading and Scene Perception*. Amsterdam: Elsevier, pp. 29–53.

Irwin, D. (2004), 'Fixation location and fixation duration as indices of cognitive processing'. In: J. M. Henderson and F. Ferreira (eds.): *The Interface of Language, Vision, and Action*. New York: The Psychology Press, pp. 105–133.

Itti, L. and C. Koch (2001), 'Computational modeling of visual attention'. *Nature Neuroscience* **2**, 1–11.

Jacob, R. J. K. and K. S. Karn (2003), 'Eye tracking in human-computer interaction and usability research: Ready to deliver the promises'. In: R. Radach, J. Hyona, and H. Deubel (eds.): *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*. Boston: North-Holland/Elsevier, pp. 573–605.

Janiszewski, C. (1998), 'The influence of display characteristics on visual exploratory search behavior'. *Journal of Consumer Research* **25**, 290–301.

Janiszewski, C. and L. Warlop (1993), 'The influence of classical conditioning procedures on subsequent attention to the conditioned brand'. *Journal of Consumer Research* **20**(2), 171–189.

Josephson, S. and M. E. Holmes (2002), 'Attention to repeated images on the world wide web: Another look at scan-path theory'. *Behavior Research Methods, Instruments and Computers* **34**(4), 539–548.

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

Karslake, J. S. (1940), 'The purdue eye-camera: A practical apparatus for studying the attention-value of advertisements'. *Journal of Applied Psychology* **24**, 417–440.

Keller, K. L. (1991), 'Memory and evaluation effects in competitive advertising environments'. *Journal of Consumer Research* **17**, 463–476.

Kent, R. (1993), 'Competitive versus noncompetitive clutter in television advertising'. *Journal of Advertising Research* **33**(2), 40–46.

Kersten, D. (1999), 'High-level vision as statistical inference'. In: M. Gazzaniga (ed.): *The New Cognitive Neurosciences*. Cambridge, MA: MIT Press, pp. 353–363.

Kersten, D., P. Mamassian, and A. Yuille (2004), 'Object perception as Bayesian inference'. *Annual Review of Psychology* **55**, 271–304.

Kieras, D. E. and D. E. Meyer (1997), 'An overview of the EPIC architecture for cognition and performance with application to human-computer interaction'. *Human-Computer Interaction* **12**, 391–438.

Koch, C. (2004), *The Quest for Consciousness: A Neurobiological Approach*. Denver (Colorado): Roberts and Company Publishers.

Koch, C. and S. Ullman (1985), 'Shifts in selective visual attention: Towards the underlying neural circuitry'. *Human Neurobiology* **4**, 219–227.

Kroeber-Riel, W. (1979), 'Activation research: Psychobiological approaches in consumer research'. *Journal of Consumer Research* **5**, 240–250.

Krugman, D. M., R. J. Fox, J. E. Fletcher, P. M. Fischer, and T. H. Rojas (1994), 'Do adolescents attend to warnings in cigarette advertising? An eye-tracking approach'. *Journal of Advertising Research* **34**, 39–52.

Krugman, H. E. (1964), 'Some applications of pupil measurement'. *Journal of Marketing Research* **1**(4), 15–19.

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

Lamberts, K. (2000), 'Information-accumulation theory of speeded categorization'. *Psychological Review* **107**, 227–260.

Laughery, K. R., S. L. Young, K. P. Vaubel, and J. W. Brelsford, Jr. (1993), 'The noticeability of warnings on alcoholic beverage containers'. *Journal of Public Policy and Marketing* **12**(1), 38–56.

Lee, T. S. and D. Mumford (2003), 'Hierarchical Bayesian inference in the visual cortex'. *Journal of the Optical Society of America* **20**(7), 1434–1448.

Leven, W. (1991), *Blickverhalten von Konsumenten: Grundlagen, Messung und Anwendung in der Werbeforschung*. Heidelberg: Physica-Verlag.

Liechty, J., R. Pieters, and M. Wedel (2003), 'Global and local covert visual attention: Evidence from a Bayesian hidden Markov model'. *Psychometrika* **68**(4), 519–541.

Loewenfeld, I. E. (1999), *The Pupil: Anatomy, Physiology, and Clinical Applications*. Boston: Butterworth Heinemann.

Loftus, G. R. and N. H. Mackworth (1978), 'Cognitive determinents of fixation location during picture viewing'. *Journal of Experimental Psychology: Human Perception and Performance* **4**, 565–572.

Lohse, G. L. (1997), 'Consumer eye movement patterns on yellow page advertising'. *Journal of Advertising* **26**, 61–73.

Lohse, G. L. and D. J. Wu (2001), 'Eye movement patterns on chinese yellow pages advertising'. *Electronic Markets* **11**(2), 87–96.

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

Matthews, G., W. Middleton, B. Gilmartin, and M. A. Bullimore (1991), 'Pupillary diameter and cognitive load'. *Journal of Psychophysiology* **5**, 265–271.

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

McFadden, S. and J. Wallman (2001), 'Shifts of attention and saccades are very similar: Are they casually linked?'. In: M. Jenkin and L. Harris (eds.): *Vision and Attention*. New York: Springer Verlag, pp. 19–39.

Mulvihill, G. (2002), 'Campbell upgrades soups to battle its competitors'. http://www.freep.com/features/food/ffill31_20021231.htm.

Musalem, A., E. T. Bradlow, and J. Raju (2008), 'Bayesian estimation of random-coefficients choice models using aggregate data'. *Journal of Applied Econometrics*, to appear.

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

Navalpakkam, V. and L. Itti (2005), 'Modeling the influence of task on attention'. *Vision Research* **45**, 205–231.

Nixon, H. K. (1924), 'Attention and interest in advertising'. *Archives of Psychology* **72**, 5–67.

Noton, D. and L. Stark (1971), 'Eye movements and visual perception'. *Scientific American* **224**, 34–43.

Öhman, A., A. Flykt, and F. Esteves (2001), 'Emotion drives attention: Detecting the snake in the grass'. *Journal of Experimental Psychology: General* **130**, 466–478.

Oliva, A. and A. Torralba (2001), 'Modeling the shape of the scene: A holistic representation of the spatial envelope'. *International Journal of Computer Vision* **42**(3), 145–175.

Palmer, S. E. (1999), *Vision Science: Photons to Phenomenology.* Cambridge, MA: A Bradford Book.

Parkhurst, D., K. Law, and E. Niebur (2002), 'Modeling the role of salience in the allocation of overt visual attention'. *Vision Research* **42**, 107–123.

Pieters, R., E. Rosbergen, and M. Wedel (1999), 'Visual attention to repeated print advertising: A test of scanpath theory'. *Journal of Marketing Research* **36**, 424–438.

Pieters, R. and L. Warlop (1999), 'Visual attention during brand choice: The impact of time pressure and task motivation'. *International Journal of Research in Marketing* **16**, 1–17.

Pieters, R., L. Warlop, and M. Wedel (2002), 'Breaking through the clutter: Benefits of advertisement originality and familiarity on brand attention and memory'. *Management Science* **48**(6), 765–781.

Pieters, R. and M. Wedel (2004), 'Attention capture and transfer in advertising: Brand, pictorial and text-size effects'. *Journal of Marketing* **68**, 36–50.

Pieters, R. and M. Wedel (2007a), 'Goal control of visual attention to advertising: The yarbus implication'. *Journal of Consumer Research* **34**, 224–233.

Pieters, R. and M. Wedel (2007b), 'Informativeness of eye-movements for visual marketing: Six cornerstones'. In: M. Wedel and R. Pieters (eds.): *Visual Marketing*. Lawrence Erlbaum, pp. 43–72.

Pieters, R., M. Wedel, and R. Batra (2008), 'The stopping power of advertising: Visual complexity, scene incongruity and message information'. Working Paper, University of Tilburg.

Pieters, R., M. Wedel, and J. Zhang (2007), 'Optimal feature advertising under competitive clutter'. *Management Science* **51**(11), 1815–1828.

Poffenberger, A. T. (1925), *Psychology in Advertising*. Chicago, Ill: A.W. Shaw Company.

Rayner, K. (1998), 'Eye movements in reading and information processing: 20 years of research'. *Psychological Bulletin* **124**(3), 372–422.

Rayner, K. and M. Castellano (2007), 'Eye movements during reading, scene perception, visual search and while looking at print advertisements'. In: M. Wedel and R. Pieters (eds.): *Visual Marketing*. Lawrence Erlbaum, pp. 9–43.

Rayner, K., B. Miller, and C. M. Rotello (2008), 'Eye movements when looking at print advertisements: The goal of the viewer matters'. *Applied Cognitive Psychology* **22**(5), 697–708.

Rayner, K., C. M. Rotello, A. J. Stewart, J. Keir, and S. A. Duffy (2001), 'Integrating text and pictorial information: Eye movements when looking at print advertisements'. *Journal of Experimental Psychology: Applied* **7**, 219–226.

Rice, B. (1974), 'Rattelsnales, french fries and pupillometric oversell'. *Psychology Today* pp. 55–59.

Rimey, R. D. and C. M. Brown (1991), 'Controlling eye movements with hidden Markov models'. *International Journal of Computer Vision* **7**(1), 47–65.

Rizzolatti, G. (1987), 'Functional organization of inferior area 6'. In: *Motor Areas of the Cerebral Cortex*. (Ciba Foundation Symposium 132). Wiley, Chichester, pp. 171–186.

Rizzolatti, G., L. Riggio, and B. M. Sheliga (1994), 'Space and selective attention'. In: C. Umiltà and M. Moskovitch (eds.): *Attention and Performance*. Cambridge, MA: MIT Press, pp. 231–265.

Rosbergen, E., R. Pieters, and M. Wedel (1997), 'Visual attention to advertising: A segment-level analysis'. *Journal of Consumer Research* **24**, 305–314.

Russo, J. E. (1978), 'Eye fixations can save the world: A critical evaluation and a comparison between eye fixations and other information processing methodologies'. In: H. K. Hunt (ed.): *Advances in Consumer Research*, Vol. 21. Ann Arbor, MI: Association for Consumer Research, pp. 561–570.

Russo, J. E. and F. Leclerc (1994), 'An eye-fixation analysis of choice processes for consumer nondurables'. *Journal of Consumer Research* **21**, 274–290.

Salvucci, D. and J. Goldberg (2000), 'Identifying fixations and sacades in eye-tracking protocols'. Eye Tracking Research and Applications Symposium 2000.

Schiller, P. (1998), 'The neural control of visually guided eye-movements'. In: J. E. Richards (ed.): *Cognitive Neuroscience of Attention*. London: Lawrence Erlbaum, pp. 3–50.

Schwartz, B. (2004), *The Paradox of Choice: Why More Is Less*. New York: HarperCollins Publishers.

Shimojo, S., C. Simion, E. Shimojo, and C. Scheier (2003), 'Gaze bias both reflects and influences preference'. *Nature Neuroscience* **6**, 1317–1322.

Singer, W. and C. M. Gray (1995), 'Visual feature integration and the temporal correlation hypothesis'. *Annual Review Neuroscience* **18**, 555–586.

Spratling, M. W. and M. H. Johnson (2004), 'A feedback model of visual attention'. *Journal of Cognitive Neuroscience* **16**(2), 219–237.

Sprott, J., J. Bolliger, and D. Mladenoff (2002), 'Self-organized criticality in forest-landscape evolution'. *Physics Letters A* **297**(3–4), 267–271.

Starch, D. (1923), *Principles of Advertising*. Chicago: A.W. Shaw Company.

Stark, L. W. and H. C. Ellis (1981), 'Scanpaths revisited: Cognitive models direct active looking'. In: D. F. Fisher et al. (ed.): *Eye Movements: Cognition and Visual Perception*. Hillsdale: Lawrence Erlbaum Associates.

Steinman, R. M. (2004), 'Gaze control under natural conditions'. In: L. M. Chalupa and J. S. Werner (eds.): *The Visual Neurosciences*. Cambridge, MA: The MIT Press, pp. 1339–1356.

Stewart, A. J., M. J. Pickering, and P. Sturt (2004), 'Using eye movements during reading as an implicit measure of the acceptability of brand extensions'. *Applied Cognitive Psychology* **18**, 697–709.

Stockman, G. and L. Shapiro (2001), *Computer Vision*. Prentice Hall.

Strong, E. K. (1920), 'Theories of selling'. *Journal of Applied Psychology* **9**, 75–86.

Tao-yi Wang, J., M. Spezio, and C. F. Camerer (2008), 'Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth-telling and deception in games'. January 17, 2008. Department of Economics, National Taiwan University, Taipei, Taiwan. http://homepage.ntu.edu.tw/~josephw/pinocchio_all.pdf, retrieved May 2008.

Teixeira, T., M. Wedel, and R. Pieters (2008), *Moment-to-Moment Optimal Branding in TV Commercials: Preventing Avoidance by Pulsing*. Working Paper, University of Michigan.

Thompson, K. G. (2005), 'Dissociation of selection from saccade programming'. In: G. R. Laurent Itti and J. K. Tsotsos (eds.): *Neurobiology of Attention*. San Diego, CA: Elsevier, pp. 124–129.

Tipper, S. P. and B. Weaver (1998), 'The medium of attention: Location-based, object-based or scene-based?'. In: R. D. Wright (ed.): *Visual Attention*. Oxford, New York: Oxford University Press, pp. 77–107.

Treisman, A. and G. Gelade (1980), 'A feature integration theory of attention'. *Cognitive Psychology* **12**, 97–136.

Treistman, J. and J. P. Gregg (1979), 'Visual, verbal, and sales responses to print ads'. *Journal of Advertising Research* **19**(4), 41–47.

Tsotsos, J. K. (1990), 'Analyzing vision at the complexity level'. *Behavioral and Brain Sciences* **13**(3), 423–445.

Ungerleider, L. G. and M. Mishkin (1982), 'Two cortical visual systems'. In: D. Ingle, R. J. W. Mansfeld, and M. S. Goodale (eds.): *The Analysis of Visual Behavior*. Cambridge, MA: MIT Press, pp. 549–586.

Van der Lans, R., F. G. M. Pieters, and M. Wedel (2008a), 'Competitive brand salience'. *Marketing Science*. forthcoming.

Van der Lans, R., F. G. M. Pieters, and M. Wedel (2008b), 'Eye movement analysis of search effectiveness'. *Journal of the American Statistical Association* **103**(482), 452–462.

Van Raaij, F. W. (1977), 'Consumer information processing for different information structures and formats'. In: W. D. Perreault, Jr. (ed.): *Advances in Consumer Research*. Vol. 4, Association for Consumer Research, pp. 176–184.

Viviani, P. (1990), 'Eye movements in visual search: Cognitive, perceptual and motor control aspects'. In: E. Kowler (ed.): *Eye Movements and their Role in Visual and Cognitive Processes*. Amsterdam: Elsevier Science Publishers, pp. 353–393.

Vogel, E. K., G. F. Woodman, and S. J. Luck (2001), 'Storage of features, conjunctions, and objects in visual working memory'. *Journal of Experimental Psychology: Human Perception and Performance* **27**(1), 92–114.

Wedel, M., W. Kamakura, N. Arora, A. Bemmaor, J. Chiang, T. Elrod, R. Johnson, P. Lenk, S. Neslin, and C. S. Poulsen (1999), 'Discrete and continuous representation of heterogeneity'. *Marketing Letters* **10**(3), 217–230.

Wedel, M. and F. G. M. Pieters (2007), 'A review of eye-tracking research in marketing'. *Review of Marketing Research* **4**, 123–147.

Wedel, M. and R. Pieters (2000), 'Eye fixations on advertisements and memory for brands: A model and findings'. *Marketing Science* **19**(4), 297–312.

Wedel, M., R. Pieters, and J. Liechty (2008), 'Temporal dynamics of scene perception: Goals influence switching between attention states'. *Journal of Experimental Psychology: Applied* **14**(2), 129–138.

Wertheimer, M. (2001), 'Laws of organization in perceptual forms'. In: S. Yantis (ed.): *Visual Perception: Essential Readings*. Psychology Press, pp. 216–224.

Witt, D. (1977), 'Emotional Advertising: The relationship between eye-movement patterns and memory — Empirical study with the eye-movement monitor'. Unpublished Ph.D. dissertation, University of Saarland, (described in Kroeber-Riel 1979).

Wolfe, J. M. and T. S. Horowitz (2004), 'What attributes guide the deployment of visual attention and how do they do it?'. *Nature Reviews Neuroscience* **5**, 1–7.

Wolfe, J. M., K. P. Yu, M. I. Stewart, A. D. Shorter, S. R. Friedman-Hill, and K. R. Cave (1990), 'Limitations on the parallel guidance of visual search: Color X color and orientation X orientation conjunctions'. *Journal of Experimental Psychology: Human Perception and Performance* **16**(4), 879–892.

Yantis, S. (2000), 'Goal-directed and stimulus-driven determinants of attentional control'. In: S. Monsell and J. Driver (eds.): *Control of Cognitive Processes; Attention and Performance*, Vol. 18. Cambridge, MA: The MIT Press, pp. 73–103.

Yantis, S. and J. Jonides (1990), 'Abrupt visual onsets and selective attention: Voluntary vs. automatic allocation'. *Journal of Experimental Psychology: Human Perception and Performance* **16**, 121–134.

Yantis, S., D. E. Meyer, and J. E. Smith (1991), 'Analysis of multinomial mixture distributions: New tests for models of cognition and action'. *Psychological Bulletin* **110**(2), 350–374.

Yarbus, A. L. (1967), *Eye Movements and Vision*. New York: Plenum Press.

Young, L. R. and D. Sheena (1975), 'Survey of eye movement recording methods'. *Behavioral Research Methods and Instrumentation* **7**, 397–429.

Zhang, J., M. Wedel, and R. Pieters (2008), *Sales Effects of Visual Attention to Feature Ads*. Working Paper, University of Maryland.