What do your eyes say? Bridging eye movements to consumer behavior

¿Qué dicen sus ojos? Conectando los movimientos oculares hacia el comportamiento del consumidor

R e v i e w

Pedro Rosa [a] [e] [*] ✉

[a] Universidade Lusófona de Humanidades e Tecnologias, Lisbon, Portugal.

[b] Cognition and People-centric Computing Laboratories COPELABS - (ULHT), Lisbon, Portugal.

[c] Instituto Universitário de Lisboa (ISCTE-IUL), Cis-IUL, Lisbon, Portugal.

[d] Centro de Investigação em Psicologia do ISMAT, Portimão, Portugal.

[e] Grupo Internacional de Investigación Neuro-Conductual, GIINCO, Barranquilla, Colombia.

ARTICLE INFO

ABSTRACT

Eye tracking (ET) is a technique that has been progressively employed to study the influence of visual stimuli on attentional processes and consumer behavior. The goals of the present theoretical article are fourfold and are based on an extensive literature revision. First, a brief historical review of ET methodology is introduced, presenting the evolution of ET techniques from the ancient proto-eye trackers to the "fresh" state-of-the-art eye ET devices. Second, the basics of ET are clarified through a simplified technical and mathematical explanation. Third, the triad eye movement-attention-consumer behavior is made clear, grounded on attention, interest, desire, and action (AIDA) theoretical model. Fourth, the most used oculometrics in marketing studies are explained and distinguished The present article addresses a number of technical and methodological issues by discussing challenges involved in ET systems and giving some guidelines for those who intend to apply ET to infer cognitive and emotional processes.

RESUMEN

El eye tracking (ET) es una técnica que se ha empleado progresivamente para estudiar la influencia de los estímulos visuales en los procesos de atención y en el comportamiento del consumidor. Los objetivos del presente artículo teórico son cuatro y se basan en una extensa revisión de la literatura. En primer lugar, se introduce una breve reseña histórica con una presentación de la evolución de los sistemas de ET. En segundo lugar, los fundamentos del ET se aclaran mediante una explicación técnica y matemática simplificada. En tercer lugar, la tríada del comportamiento, movimiento ocular y atención del consumidor se hace clara, fundamentada en el modelo teórico de la atención, interés, deseo y acción (AIDA). En cuarto lugar, se explican los indicadores oculares más utilizados en estudios de mercado. El presente artículo proporciona directrices para aquellos que tienen la intención de aplicar ET para inferir los procesos cognitivos y emocionales.

[*] **Correspondence:** Pedro Rosa, Universidade Lusófona de Humanidade e Tecnologias. Address: Campo Grande, 376 1749 - 024 Lisbon, Portugal. E mail: **pedro.rosa@ulusofona.pt**

1. INTRODUCTION

The human eye does not perceive the world with a uniform resolution, therefore eye movements (EM) serve to direct the drive the fovea (fovea centralis) which is a central pit composed of closely packed cones, in the retina, to potentially relevant locations in the visual scene (Schütz, Braun, & Gegenfurtner, 2011). This process is called foveation (See Figure 1). Greatest details (highest resolution) are found at the point of gaze and start to fade rapidly towards the periphery (Thornton & Gilden, 2007). The reader can test this assumption by looking at the period at the end of this sentence and trying to read the title at the top of this page. As a direct consequence of foveation, the human visual system uses EM combined or not with head and body movements, to direct gaze to several areas in the visual field and build a detailed map of the scene. EM are debatably the most frequent of all human movements (Bridgeman, 1992). Indeed, most human and nonhuman animals with developed visual systems actively control their gaze use eye or head movements (Land, 1999). EM are a consequence of the massive amount of visual information that bombards continuously the organism. Rather than allocate resources to processing all visual stimuli in detail, evolution appears to have selected a solution whereby small areas of the visual field are prioritized and inspected in a quick sequence (Treue, 2001). EM, therefore, are essential to the operation of the human visual system. Due to their close relation to attentional mechanisms EM can inform researchers about the affective content of visual stimuli and how stimuli are related to the individual's needs (Nummenmaa, Hyönä, Calvo, 2006; Rosa, Esteves, & Arriaga, in press). In fact, behavioral measures such as reaction times, are still used massively as a proxy of attentional processes, however the use of key-press responses may reflect more than attentional processes (Armstrong & Olatunji, 2012). Furthermore, self-report instruments, used to infer cognitive and attentional processes, are highly susceptible to exogenous influences, sustaining the use of complementary measurements (Schwarz, 1999). To overcome the above limitations, eye movement recording or ET (ET) can be a non-intrusive approach (Rosa, Esteves, & Patrícia, 2014). As EM can be controlled volitionally (over attention), at least to some degree, EM recording can be a rich source of information about human attentional selection. Therefore, ET has been successfully applied in Psychological (e.g., Calvo & Lang, 2004) and medical domains (e.g., Lahiri, Warren, & Sarkar, 2011) but as

well as ergonomics, advertising and design, allowing researchers to obtain measures closely related to attention and emotion (Rosa, Esteves, & Patrícia, 2010, 2012, 2014; Wedel & Pieters, 2006). Until recently, the use of ET system was somewhat limited, because its use was complex (e.g., manual calibration, image optimization), time consuming for researchers and participants and quite expensive. Yet, recent commercial generation of ET systems, called in this article as "ET 2.0" (in analogy with web 2.0), use relatively low-cost cameras, allowing a quick calibration an accurate estimation of the point-of-regard are still unknown (e.g. The eye tribe),.Parallel with this increase in availability of ET 2.0 systems, there is a trend to develop software for scientific ET on an open-source basis, making it easily and freely available to all researchers. (Dalmaijer, 2014). ET. 2.0 systems due to its mobility and flexibility, allow researchers to record EM in non-laboratory contexts. In fact, researchers are able now, to set up a quick and portable eye analysis laboratory wherever participants are (Rosa et al., 2014). Such kind of tools applied in psychological research should be thought as an open door for further research on visual marketing and might help to answer to the following question: "Did I decide to buy this? "

We tend to say to ourselves "yes", but the question goes further. Damásio (2003) raise the following question "Is the information I received through the senses responsible for the perception of what I see? In fact, the sensory world, held by the senses is a creation of the brain. What we see is not a precise copy of what "is there", but a creation of reality elaborated by the brain (Morin, 1987). But, where do our eyes move when we are faced with a store full of many colorful displays, offers or specific seasonal attractions? Is the attention drawn by one type of package over another in this situation? ET may aid to answer these questions. ET allows us knowing on where people are looking, for instance, which area of the product, which part of the shop, which shelve, which promotional stands and offer boards and which packaging. Commonly, market researchers intend to "offer" product information to consumers in an effective way, so that consumers can be aware of the existence of a given product. If consumers identify a product that likely satisfies their needs, it is thought that consumers tend to purchase this specific product again. Oculometric data provided by ET can be an important component of marketing, by orienting the creation of a better brand logo, packaging, advertising design, and even more recently, web page design (e.g., Duchowski, 2007; Goldberg & Wichansky, 2003; Wedel & Pieters, 2006).

Figure 1. *(a) Fovea centralis in retina. Image retrieved 1 June, 2015, from https://ece.uwaterloo.ca/~z70wang/research/fovea/fovea_model.html (b) Simulation of foveation process in the human eye. Image retrieved 1 June and from http://www.utdallas.edu/~tres/integ/sen3/display7_11.html.*
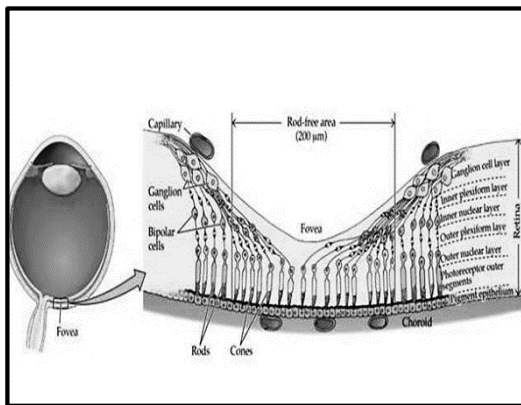
(a)

(b)

Knowing exactly what captures consumers' attention, and what does not, in specific selling scenarios, ET offers the possibility to get new insights and can be a unique advantage in consumer decision-making. In fact, as mentioned by Wedel and Pieters (2006), if "seeing is believing" than "believing is buying", so, eye movement recording can aid managing what consumers seek to maximize profits. The enthusiasm for using ET in consumer science relies on the desire to understand how consumers choose their products and to make the most of the product and its context to capture and locate consumers' attention on it. Since ET 2.0 systems are still recent, this article relies on an in-depth review on ET system and methods and is thought to be a guidance document for researchers who feel somehow "seduced" to start using these kind of research tools. It addresses a number of technical issues by discussing challenges involved in using ET 2.0 which is low-cost and non-obtrusive, but not always user-friendly and accurate.

2. METHOD

The present study is based on a systematic review that explores the use of eye tracking in consumer science. A comprehensive scientific literature search was performed in scientific databases such as The Library of congress, PubMed and EBSCO for "consumer science" or "marketing" intersected with eye, tracking, ET, EMs, gaze, saccade, fixation or pupil. The authors of studies revealed by this search were

then entered as search terms into the same databases, in order to uncover additional eye tracking studies. There were three primary criteria for inclusion of studies in the review, in order to ensure commensurability: 1) Studies that used eye tracking, independently of their technique; 2) Attentional/emotional processes had to be studied 3) Non-clinical context (healthy participants). Auto- and manual-search for duplicates was performed through ENDNOTE X7. All duplicates were removed, after being identified. At the end, only 58 valid scientific articles/books from 1998 to 2015 remained. However, and since this article intend to create an historical contextualization of ET usage, more 26 references were added to this list, mostly old ET referential books, yet relevant. Altogether, 84 articles and books provided scientific information that addressed the purpose of the present article.

### 2.1 Eye movement recording: a little splash of history

The first reports of EM date back to the 18th century (Porterfield, 1737). In the 19th century, the French ophthalmologist Louis Javal (1879) observed the EM during reading were not smooth, but rather a combination of fixations (the maintaining of the visual gaze on a spatial location) and ballistic EM, coined by him as ocular saccades. Ahrens (1891), along with Delabarre (1898) and Huey (1898) were the firsts to attempt recording EM through a sort of contact lens with a hole for the pupil. The lens was coupled to an aluminum lever that moved in response to the ocular

movement. Despite this painful system, Huey showed the existence of regressive saccades (EM to the previous regions of the text). Some years later, Dodge and Cline (1901) were the first to study EM with a non-intrusive photographic methodology and light reflections. Based on motion picture photography, Judd and colleagues (1905) carried out the first frame-by-frame analysis and empowered quantitative research on a solid basis. The photographic methods were less obtrusive than early mechanical methods and more precise, however, they were tiresome and lengthy in terms of analyzing data. Along with Judd and colleagues, Buswell (1935) studied and analyzed EM in reading within different ages and levels of schooling, leading to several leaps in different scientific domains. Unlike earlier devices used in early reading studies, Buswell's eye tracker was specifically made for the purpose of recording both the horizontal and vertical EM.

After the second world war, an extensive eye movement recording was conducted by in Fitts and coworkers (1950), publishing several technical reports that encompass over 500,000 video frames of over 40 pilots taken under several flight conditions. The results of this study allowed a more efficient arrangement of instruments and identified which were difficult to read, for a possible redesign of the cockpit. In the 1950s, Yarbus (1967) conducted one of the most referenced studies in ET. In his studies, observers' EM were recorded while viewing natural objects and scenes (see Figure 2). He showed that point-of-gaze tended to converge in relevant areas in the visual field. In the 1970s, much of the progress in ET was focused on technical improvements to increase accuracy and precision, and reduce the noise (the impact of the eye tracker on the observer). In this time, ET research grew quickly, particularly in reading research (For an extended review, see Rayner, 1998).

Figure 2. *Examining Ilya Repin's painting (the unexpected visitor) with different instructions. (a) Free viewing. (b) Estimate the material circumstances of the family. (c) Estimate the ages of the people. From: Duchowski, A. (2007). Eye tracking Methodology: Theory and Practice.*
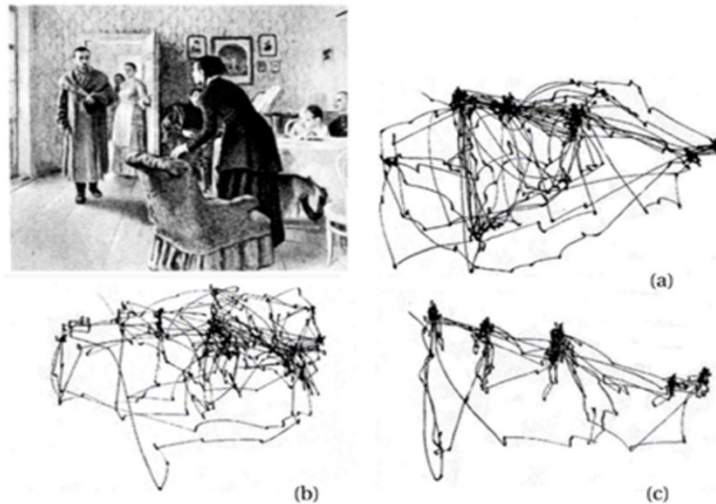
In the early 1980s, a strong influential eye-mind hypothesis was formulated by Just and Carpenter (1976). They found that readers fixate words in a different way, depending on its content (Just & Carpenter, 1980). These findings revealed that not every word in a text is fixated, but there are syntactic and semantic components of each word that determines if an ocular fixation occurs or not. Their research supported the "eye-mind" assumption, that is, when a subject looks at a word or an object (fixating it) he/she also "thinks" about it (cognitive processing). This assumption is frequently taken for granted by earlier eye tracker researchers.

Since the 80's, personal computers have become powerful enough to do real-time eye movement recording, becoming possible the use of video-based eye trackers. Due to Moore's Law (1965), the processing speed, memory capacity, the number

and size of pixels in digital cameras and improved computer vision techniques, ET manufacturers have developed the fourth generation and state-of-the-art "easy-to-use" eye trackers (Duchowski, 2007).

Since the 1990s up to now, the dropping prices of tracking systems have increased the use of the eye trackers, being used within a wide spectrum of scientific domains, especially in marketing research (Gamito & Rosa, 2014; Wedel & Pieters, 2006). In fact, a novel generation of ET systems became more attractive as they use relatively low-cost cameras, are lesser intrusive, allowing a quick calibration process with motion correction algorithms for large head movement and an accurate estimation of the point-of-regard. (e.g., The Eye Tribe). These kind of recent ET systems can be thought as ET 2.0 in analogy with web 2.0 (Rosa et al., 2014).

2.2 The basics of Eye Tracking

The eyes and their movements are described as a window to the mind. They play a vital role in expressing a person's desires, cognitive processes and emotional states (Underwood & Foulsham, 2006). The importance of EM is implicitly related to how we gather essential information to identify the properties of the visual world. If we track eye movement with high-levels of accuracy and precision, we are able to know which areas of the visual space were minutely processed, that is, where visual attention was located (IBanović, Rosa & Gamito, 2014; Duchowski, 2007). Since EM are indexes of the covert visual attention (Posner, 1980), ET methodologies can offer new insights into moment-to-moment visual stimuli processing (e.g., print advertising, ad placement, web pages, packaging) and consumer decision (Wedel & Pieters, 2006).

According to Duchwosky (2007), there four common techniques to record EM (1) Electro-Oculograph (EOG) based, which measures electric potential differences when the eyes are moving. In this technique, the eye movement potentials range between 15–200 µV, with a nominal sensitivity around 20 µV/deg of eye movement.; (2) Scleral contact lens (SCL) based methods, in which a mechanical or optical reference object (e.g. reflecting phosphors, wire coils) mounted on a contact lens, enabling a precise eye movement measurement (Young & Sheena, 1975); (3) Photo- or Video-OculoGraphy and (POG or VOG) based method, in which some particular eye measurements can be made (e.g. shape of pupil, position of limbus). However, mentioned eye measurements may not be

made in an automatic fashion, involving in fact visual examination of recorded EM on videotape. This means that the eye movement assessment is performed through an unexciting and likely to error frame-by-frame analysis (4) Video-Based infrared oculography (VIROG), which is the leading method nowadays. VIROG technique uses relatively low-cost cameras and image processing hardware to estimate the point-of-regard (POR) or point-of-gaze, that is, the (x,y) coordinates of the user's gaze on the displayed stimulus, in real-time (Holmqvist et al., 2011). These systems may be table-mounted or head-mounted. In fact, the latter method, named also infra-red (IR) ET due to use of IR light to estimate user's POR, is the most popular method for current ET-based market research. VIROG technique has captured attention of several high-profile technology companies. This is also reflected in established and new manufactures of eye trackers that have started to produce ET 2.0 systems, such as GazePoint's GP3 tracker ($495), Tobii's EyeX tracker ($139), and the EyeTribe tracker ($99) (Dalmaijer, 2014).

A profound description of ET methodologies is provided in Duchowski (2007). The corneal reflection of a given light source (usually IR) is measured and estimated relative to the center of the pupil. These light reflections on the cornea are known as the Purkinje Reflections (Crane, 1994). Since Purkinje reflections are formed, VIROG systems identify the first Purkinje image trough calibration procedures. (See Figure 3). Any free or commercial eye-trackers identify and measure only the first Purkinje reflection (glint) off the front of the cornea (in a monocular or binocular mode), providing appropriate accuracy and precision. Most VIROG systems use one camera and one or two infrared light sources that need usually 5 or 9, point calibration procedure (e.g., GazePoint's GP3 tracker or the EyeTribe tracker). The difference between the pupil center and the corneal reflection might change with eye rotation, but this difference is quite constant while the head is moving (Mele & Federici, 2012). Assuming a fixed reference point (glint) on the user's eye and that the user's position is stable, is possible to calculate the POR. The ET can be made from the user's eye image by integrating the information related to the center of the pupil and the position of the reference point.

Figure 3. *(a) The four Purkinje images. Image retrieved 14 November, 2014, from http://www.howtodegrid.com/dconcept.htm (b) Relative positions of pupil and first Purkinje images in eye tracker camera. From: Duchowski, A. (2007). Eye tracking Methodology: Theory and Practice.*
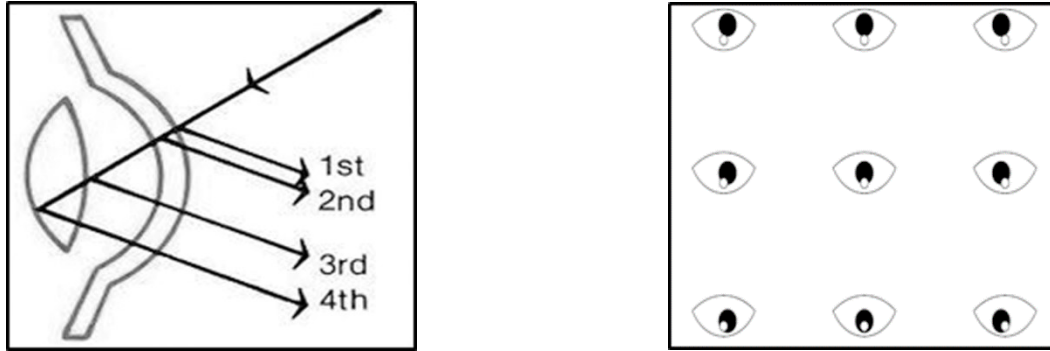
(a)

(b)

Having these two pieces of information, the position of the pupil center relative the fixed reference point is calculated, obtaining the distance in a coordinate system for the image. This distance ($S = (x_s, y_s)^t$) is converted to a coordinate system on the computer screen ($S = (x_s, y_s)^t$) through two second-order equations:

$$x_e = a_0 + a_1 x_s + a_2 y_s + a_3 x_s y_s + a_4 x^2 + a_5^2$$

$$y_e = a_6 + a_7 x_s + a_8 y_s + a_9 x_s y_s + a_{10} x^2 + a_{11}^2$$

Where a_i, $i = 0;.., 11$ are the coefficients of the quadratic polynomials. These coefficients are achieved through a calibration procedure mentioned before, obtaining the distance values $E_i = (x_{ei}, y_{ei})^t$ for nine distinct points $S_i = (x_{si}; y_{si})^t$ on screen ($i = 1,.., 9$), which allows the resolution of the two overdetermined systems (6*9), one for the first equation and another one for the second equation.

However, the methods used for gaze estimation may vary significantly across different eye-tracking systems, depending on the illumination source. (e.g., Bhaskar, Foo, Ranganath, & Venkatesh , 2003; Zhiwei & Qiang, 2007). Bright or dark pupil tracking techniques are based on the position of illumination source. If the lighting source is in line with the optical axis, causes a reflection of the retina, then the bright pupil effect occurs. This is quite similar to what happens in flash photography to create a red-eye effect. On the other hand, a dark pupil tracking technique occurs in a photograph or video image when the lighting source is

offset from the optical axis, and the light reflection is projected away from the camera, making the pupil look black. Each technique has its own benefits and drawbacks, depending on the conditions under which technique is applied and the physical characteristics of the user, whose EM are being tracked (Duchowski, 2007). However, and according to Holmqvist and colleagues (2011), the comparison between oculometrics from different manufacturers' systems involve always some risk, as the pupil tracking techniques and the algorithms used by manufacturers may significantly differ.

2.3 The intimacy between visual attention and EM: the central piece for using ET in consumer science

The idea of selectivity of attention was initially supported by William James (1890), and well-exemplified in the Latin phrase 'Pluribus intentus, minor est ad singular sensus. He tried with this phrase to state that humans are limited and not attend to all things at once. In fact, attention is the selective process that coordinates the perception-action cycle and preserves goals over time, despite its limited capacity. The brain processes visual sensory information by focusing on specific visual elements of sensory "bombing", being some of those examined in more detail (high resolution) through the foveation process. The inspection of the visual field is minutatim performed, but not in its entirety. This imperious information selection characterizes the visual selective attention.

In any visual stimuli (e.g., chocolate bar), attention can be deployed in one of two ways:

endogenously or exogenously (Posner, 1980). In endogenous attention, attention is assumed to be under the overt control of the subject, (e.g., "I am hungry, and I will attend to the food shelf"). This is also thought as "top-down" or goal-driven attention (Rosa, et al., 2012; Yantis, 1998). Endogenous attention is voluntary, but has a slow time course. In contrast, attention can be also reflexive or exogenous when is driven by an external stimulus, that automatically captures attention to a specific visual area. This has been defined as "bottom-up" or stimulus-driven attention. For instance, a red chocolate package among non-colorful chocolate packages will capture attention exogenously. Exogenous attention attracts attention in an automatic fashion and has a fast time course (Cheal & Lyon, 1991). When a specific visual area or object in a scene is attentively selected, means that it will be processed at high resolution and other visual areas or elements are concurrently suppressed. This attentional process is achieved through bottom-up and top-down mechanisms (Rosa, et al., 2014). The first is related to the visual elements (e.g., contrast, luminance), the latter is initiated from higher cortical centers and driven by affective states, goals, memory or context (Rayner & Castelhano, 2007).

The combination of these mechanisms, along with other cognitive faculties is the base of the attentional theory to visual marketing or product communication, as depicted in Figure 4 (For more detail see Wedel & Pieters, 2006). According to this attentional theory, a particular product (e.g., Colgate toothpaste) in a cosmos of clutter (large supermarket shelves), can grab attention depending on its informativeness (goal-relevance) and salience (contrast), and this can be reflected on EM (Yantis & Jonides, 1990).

When a visual stimulus/product is presented (advertisement, billboard or website), specific features (singletons) are automatically processed. In fact, these basic perceptual features can be pre-attentively processed (e.g., Öhman, Flykt, & Esteves, 2001; Rosa, Gamito, Oliveira, & Morais, 2011; Rosa, Gamito, Oliveira & Morais, 2011, Rosa et al., 2014). Also, when someone goes in a store, creates automatically a conspicuity map. This map is nothing more than a topographic map, coding all the relevant visual areas of a particular scene for orienting attention to visual spots of interest (Koch & Ullman, 1985).

Already in 1897, Harlow Gale found similar findings, supporting the idea that different colors used in advertisements influenced readers' attention (Spielberger, 2004). These evidences support also the attention, interest, desire, and action (AIDA) theoretical

model based on Scott's work (1908). According to this model, in the first stage (attention), any given advertising stimulus must first draw attention of consumers and at the final stage (action), consumers take action, i.e., they decide to buy a product. Indeed, marketing techniques can affect attention and improve advertising effectiveness in several ways (e.g., Coulson, 2000; Ferreira et al., 2011; Warde, 1999). It is known that pictures are a useful tool as attention grabbers in visual marketing. For example, extensive use of sex appeal and celebrity endorsement in advertising can be understood in this context. Based on Paivio's work (1991), images and text seemed to be processed differently, whereas pictures are processed more holistically and verbal information is thought to be processed more sequentially (Kahneman & Henik 1981).

Hence, the two-thirds of print advertisements today have pictures covering more than half of the existing spatial area is not surprising (Chandon, Hutchinson, Bradlow, & Young, 2009).

Additionally, attention occurs by either overtly and covertly (Posner, 1980). Overt attention is achieved when the gaze is pointed towards a selected location using a directly observable combination of eye, head and body movements. Instead, covert attention, usually designated as "looking out of the corner of the eye", occurs when attention is oriented to a specific location in the absence of directly observable eye movement. In fact, when looking at a complex visual scene, is possible to dissociate covert and overt attention, however, that takes some of our conscious efforts (e.g., fixating a point in the space, but the attention is kept elsewhere). Since there is substantial evidence that covert attention precedes EM to an anticipated target, both in response to endogenous or exogenous signals (Posner, 1980), ET allows exploring the relation between visual marketing stimuli (e.g., print advertisements, television commercials), EM, perception, attention and decision (e.g., Afonso, Colaço, Sargento, & Rosa, 2011). Therefore, EM can be a comprehensive index of cognitive information processing for different stimuli (e.g., tv commercials, print advertisement, web advertisement, packaging). In fact, EM can discriminate whether a particular characteristic of an advertisement (static or video) was noticed by the user or not, improving marketing strategies and its impact on potential consumers. Additionally, EM can be a reliable correlate of emotional information even when visual cues cannot be consciously perceived (e.g., Rosa et al., 2014; Rosa & Morais, 2011). In this manner, it is possible to tailor and

test a product in such way that the consumer´s attention is delivered in a fast and efficient mode.

Figure 4. *Wedel and Pieters's adapted schematic diagram for the representative attentional factors in visual marketing. From: Wedel, M., & Pieters, R. (2006). Eye Tracking for visual marketing.*
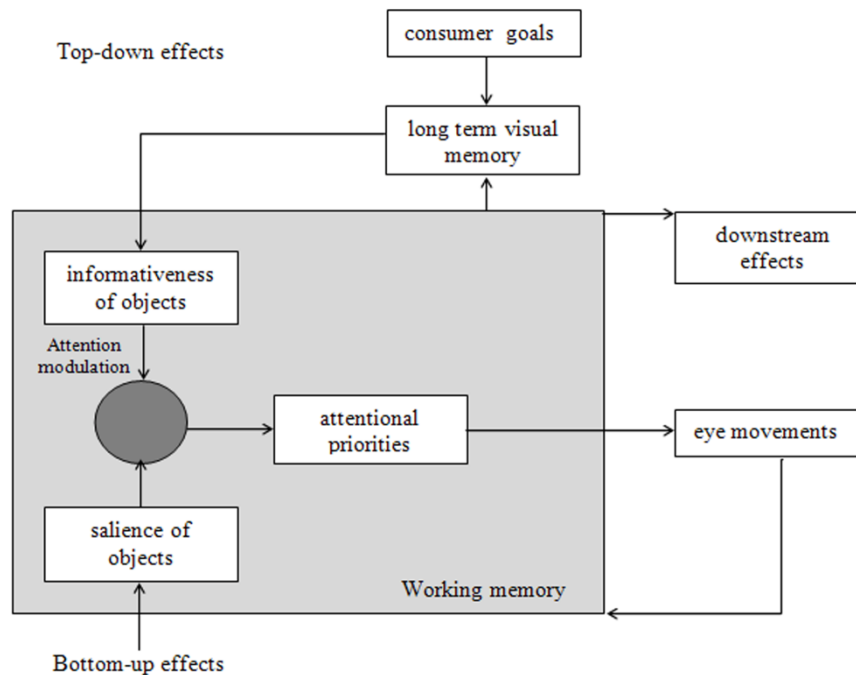
2.4 Most used oculometrics and its inferences on consumer behavior and marketing research

Both humans and animals show a consistent pattern of EM which can be defined as a "saccade and fixate" strategy (Land, 1999). When someone looks at a scene or search for a specific area in the a visual field, his/her eyes move every 250–350 ms. These movements serve to move the fovea centralis (the high resolution area of the retina) to the part of the visual field that is going to be processed in detail (Rayner & Castelhano, 2007). Around 90% of viewing time is spent in ocular fixations, in which gaze is held almost stationary. When visual attention is attracted to a new area in the visual field, ballistic EM (saccades) are made, positioning the fovea on the new area for detailed processing (Land & Hayhoe, 2001). During ocular fixations information is taken in, while during saccades no information processing is done. In matter of fact, during ocular saccades, we are effectively blind.

There are two motives for this ocular strategy. Firstly, the fovea is fairly a small area. Depending on exactly how it is defined, its angular diameter ranges between 0.3º and 2º, and the foveal depression covers only about 1/4000th of the retinal surface (Steinman, 2003). Secondly, in order to get visual information, eye gaze should be still between saccades, therefore, the photoreception is relatively slow (around 20 ms) (Friedburg et al., 2004).

Most studies with eye movement measurement is based on these two EM, which are extra-ocular movements. Fixations in marketing research should be interpreted with some precautions, since depends on the context. In fact, the majority of the reviewed articles/books was focused on ocular fixations (73.3%; n=65). For instance, in an encoding task (e.g., browsing a web page), higher fixation frequency in a particular spatial area (previously defined or not by the researcher) might be an index of greater interest in the

target, such as a picture on a website page, or it can be a sign that the target is complex in some way and more difficult to encode (Just & Carpenter, 1976). However, these interpretations may be upturned when a search task is presented: A higher number of fixations is usually an index of greater uncertainty in identifying a target (Jacob & Karn, 2003). Regarding the duration of a fixation, is generally accepted that is linked to the processing-time applied to the object/product/stimulus being fixated (Just & Carpenter, 1976).

Regarding ocular saccades, about one-third of the reviewed articles/books were related to saccadic movements (n=24). An ocular saccade occurs when both eyes are moving in the same direction between two fixations. There is no encoding information during this ballistic movement. Inferences about salience or perceptive complexity are quite hard to make. Although, backtracking EM (regressive saccades) can be thought as a processing difficulty index (Rayner & Pollatsek, 1989), but can equally be used as a measure of recognition, as is an inverse measure between the number of regressions and the salience of a given sentence.

Only 8,4% of the reviewed articles/books was related to scan path analysis, that is, a complete saccade-fixation-saccade sequence analysis. ET can aid identifying optimal scan paths, for instance, during a visual search task. An optimal scan path is represented by a straight line to a desired target, with a relatively short fixation duration (Goldberg & Kotval, 1999).

Regarding to pupil reactivity, which is an intra-ocular movement, only a few articles/books stressed out this measure (< 5%; n=4). However, pupil dilation (change in pupil size), that is an important index of cognitive load and emotional state (For an extensive review see Holmqvist et al., 2011). For single pupillary responses, dilation occurs 2-7 seconds after emotional stimuli are presented and faster dilation (time to peak) generally happens for stronger emotional stimuli (Hess, 1972). Partala and Surakka (2003) revealed that there was no response for the initial 400ms, and then a steep increase in pupil diameter peaking at 2-3 seconds after stimulus onset, using different emotional auditory stimuli. The emotional content in the visual stimulus might trigger a proportional pupil reaction as documented by Steinhauer and colleagues (1983) that showed that highly aversive or pleasant pictures were linked to larger dilations. Given that pupil dilates with sympathetic activity and constricts with parasympathetic activity (Steinhauer, Siegle, Condray, & Pless, 2004), pupil size can be an accurate and non-intrusive measure of emotional arousal (e.g., Partala &

Surakka, 2003; Rosa et al., 2010). Regardless of considerable development in ET technology (e.g. ease of use, improved accuracy, and enhanced sampling rate), accurate pupillary measurement depends on several factors such as participant's gaze angle or illumination issues (Porter, Troscianko, & Gilchrist, 2007) If emotions and pupil size variation are reliably associated with each other, then pupil tracking offers a possibility for unobtrusive monitoring of emotion-related reactions during the exposure to a visual stimulus, because no attached sensors are needed. However, pupil size is determined by many other factors, such as luminance, chrominance or contrast, so it is a very sensitive measure, likely to contamination (Goldberg & Wichansky, 2003; Rosa, Caires, Costa, Rodelo, & Pinto, 2014).

In fact, most methods in marketing rely on indirect measures, such as the subject's self-report. With a true direct measure of importance, such as pupil size and previous mentioned oculometrics, researchers/marketers are able to explain the impact of visual information, one of the areas where consumer behavior studies should have important inputs for marketers and policy-makers.

2.5 Eye Tracking is not a bed of roses: some limitations to be aware of.

There is not only advantages of using ET in consumer science. In fact, EM measurements can be misleading. Therefore is crucial for researchers to have an advanced knowledge of basic cognitive processes such as perception and attention. ET records and displays ocular fixations, but does not record any measure related to peripheral vision, which makes up 98% of the human visual field (Land, 1999). However, the rationale of using ET is that covert attention is followed by overt attention, because peripheral vision allows to choose where to fixate our fovea next. By other words, we can see (process) some specific visual element in our visual field without directly fixating on it. For instance, banner ads can be seen on the right side of a page, using peripheral vision. Information from their position, visual appearance or from previous web experience, is gathered and used to identify them as ads, and potential consumers frequently choose not to fixate on them. For instance, in a heat map, some researchers may assume that users only see the fixated areas, or hot spots (Figure 5b). It is important to recognize that the ocular fixations do not represent everything that was seen by the consumers/users (Figure 5c). They are just the areas where consumers/users fixated on, viewing them in detail/high resolution (Figure 5b)
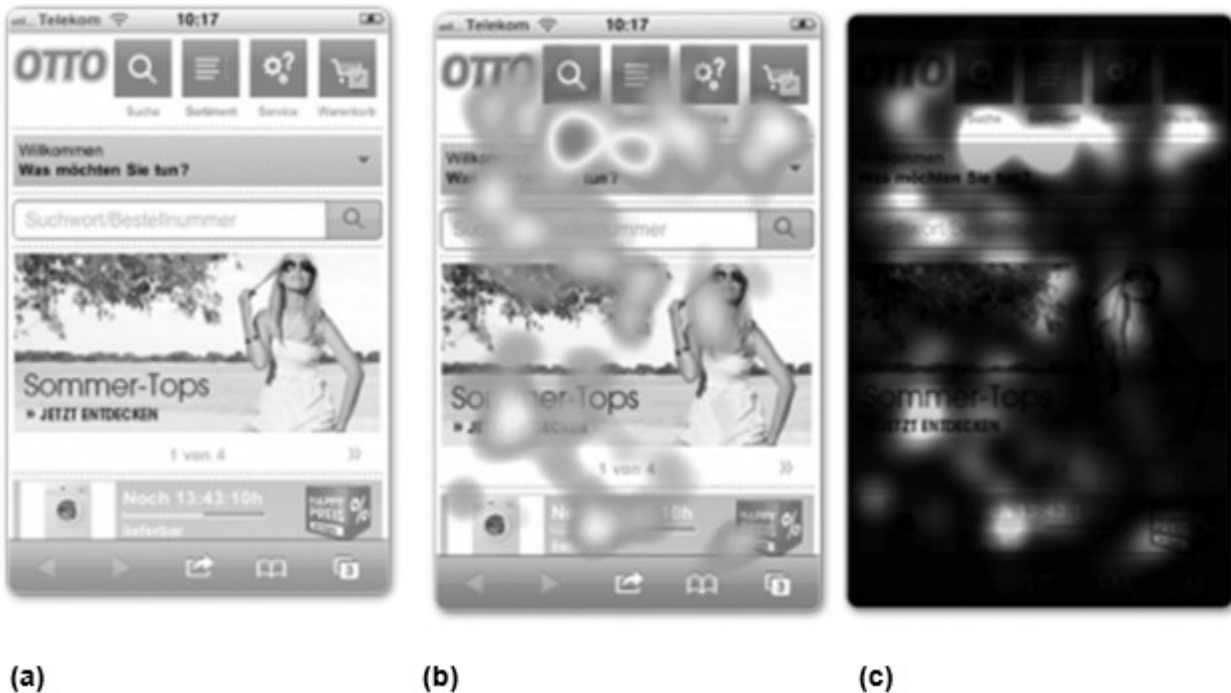
Figure 5 shows two versions of different maps. A gaze opacity map can be though as the reverse of a heat map because only fixated areas are shown and the remained areas are masked or blacked out. It is quite obvious that consumers/users do not process web pages like this. User process visual information through foveal and peripheral vision. It is crucial that researchers who had intentions to use ET, independently of the scientific domain, should be aware of this.

In order to get accurate eye measurements, users should perform the experimental tasks without interruption. Therefore, asking users to think aloud, asking probing questions, or even allowing participants to ask questions interferes with data accuracy (Hyrskykari, Ovaska, Majaranta, Räihä, & Lehtinen, 2008). Instead of using a think-aloud technique, researchers should rather apply a retrospective think-aloud technique, in which participants are asked to describe what they were doing while a recording task or tasks they have performed is played back. Nonetheless, when using this technique, user's comments may lose some of their immediacy (Elling, Lentz, & De Jong, 2011).

Despite being unobtrusive technique ET systems can interfere with the naturalness of the tasks. The perfect ET system should be completely unobtrusive to users, however, using head-mounted systems, ET glasses or even sometimes a chin-rest may create artificial situations with low ecological validity, being more difficult to generalize results to real contexts (Johansen & Hansen, 2006).

Furthermore, and regardless of the availability of ET 2.0 systems and free software for ET and EM analysis [e.g., OpenGazeAndMouseAnalyze software (OGAMA: Voßkühler, Nordmeier, Kuchinke, & Jacobs 2009), planning and setting up an ET study can add some extra time the whole experimental process. ET also generates big eye data, adding substantial time to the data analysis process. Finally, the more technology a researcher has, the more can go wrong. Despite of their reliability and low-cost, recent ET systems, and especially ET 2.0 systems, can be more demanding and require more technical skills. Therefore, researchers cannot expect an "easy game" and should be prepared for some troubleshooting.

Figure 5 *(a) An example of a visual content on a mobile device. (b) Heat map c) Gaze opacity map Retrivied from: http://www.uxmatters.com/mt/archives/2009/10/ ~eyetracking-is-it-worth-it.php*

(a)   (b)   (c)

3. CONCLUSION

This paper is the result of an historical and technical review of ET techniques and a brief guide to those who are interested to use ET systems, not only in consumer science, but also in other scientific domains. The present systematic review, allowed us to know that ET has a history of 100 years of research, but is not yet been standardized. The use of EM to infer users' cognitive and emotional processes was demonstrated and they can, in fact, to contribute to the understanding of consumers needs and desires. Therefore, ET techniques can afford useful advantages for the in-depth analysis of visual marketing stimuli, however, limitations should be always considered and researchers should be aware whether ET techniques will contribute or not to what they intend to test. It is always hard to answer to the question "when should I use ET?" Then, this article should serve as a guidance document and orient everyone that expects to use ET in consumer research. There are success cases in the literature in which EM have enriched our understanding of a particular phenomenon and it is the ultimate appeal of the ET. By the use of ET methods it is possible to capture consumer behavior in real-time. Thus, other EM such as eye blinks (periocular movements), vergence or accommodation movements can be explored and they may also be a rich source of data. Future research is encouraged to extend the measures reported here to the examination of how other EM can be used in consumer science.

4. REFERENCES

Afonso, R., Colaço, N, Sargento, P., & Rosa, P. J. (2011). Processo de Tomada de Decisão no Feminino: O ciclo menstrual como mediador. In F. Rodrigues, *A influência do Neuromarketing e a tomada decisão* (pp. 117-137). Viseu: Psicosoma.

Ahrens, A. (1891). *Die Bewegung der Augen beim Schreiben*. Rostock: University of Rostock.

Armstrong, T., & Olatunji, B. O. (2012). Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis. *Clinical Psychology Review*, 32, 704-723.

Bhaskar, T., Foo Tun K., Ranganath, S., & Venkatesh, Y. (2003). Blink detection and Eye Tracking for eye localization. *Conference on Convergent Technologies for Asia-Pacific Region*:

TENCON'03; 15–17 Oct 2003; Bangalore, IN; 2003. pp 821–824.

Bridgeman, B. (1992). Conscious vs unconscious processes: The case of vision. *Theory and Psychology, 2*(1), 73-88

Buswell, G. T. (1935). How people look at pictures. Chicago: University of Chicago Press.

Calvo, M., & Lang, P. (2004). Gaze Patterns When Looking at Emotional Pictures: Motivationally Biased Attention. *Motivation and Emotion, 28*(3), 221–243.

Chandon, P., Hutchinson, J., Bradlow, E., & Young, S. (2009). Does In-Store Marketing Work? Effects of the Number and Position of Shelf Facings on Brand Attention and Evaluation at the Point of Purchase. *Journal of Marketing, 73*(6), 1-17.

Cheal, M., & Lyon, D. R. (1991). Central and peripheral precueing of forced-choice discrimination. Quarterly *Journal of Experimental Psychology*, 43A, 859–880.

Coulson, N. (2000). An application of the stages of change model to consumer use of food labels. *British Food Journal*, 102, 9, 661-8.

Crane, H. D. (1994). The Purkinje Image Eyetracker, Image Stabilization, and Related Forms of Stimulus Manipulation. In D. H. Kelly (Ed.), *Visual Science and Engineering: Models and Applications* (pp. 13–89). New York: Marcel Dekker.

Dalmaijer E. (2014) Is the low-cost EyeTribe eye tracker any good for research? *PeerJ PrePrints,* 2, 1-35. Retrieved from https://peerj.com/preprints/585v1.pdf

Damásio, A. (2003). *Looking for Spinoza:* Joy, Sorrow, and the Feeling Brain. Harcourt. Orlando, FL.

Delabarre, E.B. (1898) A method for recording eye-movements. *The American Journal of Psychology, 9*, 572–574.

Dodge, R., & Cline, T.S. (1901). The angle velocity of eye movements. *Psychological Review*, *8*, 145–157.

Duchowski, A. (2007). *Eye Tracking Methodology: Theory and Practice* (2nd ed.). London: Springer-Verlag.

Elling, S., Lentz, L., & De Jong, M. (2011). Retrospective think-aloud method: Using eye movements as an extra cue for participants' verbalizations. Proceedings of the 2011 *Annual Conference on Human Factors in Computing Systems*, 1161- 1170. New York, USA: ACM

Ferreira, P., Rita, P., Morais, D, Rosa, P, Oliveira, J., Gamito, P., Santos, N., Soares, F., & Sottomayor, C., (2011). Grabbing attention while reading website pages: the influence of verbal emotional cues in advertising. Journal of *Eye Tracking, Visual Cognition and Emotion*, *1*(1), 64-68.

Fitts, P., Jones, R., & Milton, J., (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeroautical Engineering Review*, 9(2), 24-29.

Friedburg, C., Allen, C.P., Mason, P.J., Lamb, T.D., 2004. Contribution of con photoreceptors and post-receptoral mechanisms to the human photopic electroretinogram. *Journal of Physiology*, 556, 819–843.

Gamito, P. & Rosa, P.J. (2014*). I see me, you see me: inferring cognitive and emotional processes from gazing behavior.* Newcastle upon Tyne: Cambridge Scholars Publishing.

Goldberg, H. J., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics*, 24, 631-645.

Goldberg, H. J., & Wichansky, A. M. (2003). Eye Tracking in usability evaluation: A practitioner's guide. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 493-516). Amsterdam: Elsevier.

Hess, E. H. (1972). Pupillometrics: A method of studying mental, emotional, and sensory processes. In N. S. Greenfield, & R. A. Sternbach (Eds.), *Handbook of Psychophysiology* (pp. 491-531). New York: Holt, Rinehart & Winston.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye Tracking: a comprehensive guide to methods and measures*. Oxford, UK: Oxford University Press.

Huey, E. B. (1898). Preliminary experiments in the physiology and psychology of reading. *American Journal of Psychology, 9*, 575-586.

Hyrskykari, A., Ovaska, S., Majaranta, P., Räihä, K.-J. & Lehtinen, M. (2008). Gaze Path Stimulation in Retrospective Think-Aloud. *Journal of Eye Movement Research, 2*(4), 1-18.

IBanović, M., Rosa, P., J., & Gamito, P. (2014). Eye of the beholder: Visual search, attention and product choice. In P. Gamito and P.J. Rosa, *I see me, you see me: inferring cognitive and emotional processes from gazing behavior* (pp.

40-61). Newcastle upon Tyne: Cambridge Scholars Publishing.

Jacob, R. J. K., & Karn, K. S. (2003). Eye Tracking in Human-Computer Interaction and usability research: Ready to deliver the promises, In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 573-605). Amsterdam: Elsevier.

James, W. (1890). *The principles of Psychology* (Vol 1). New York, NY: Henry Holt & Co.

Javal, L.(1878). Essai sur la physiologie de la lecture. *Annales d'Oculistique, 79,* 97-117.

Johansen, S. A., & Hansen, J. P. (2006). *Do We Need Eye Trackers to Tell Where People Look? In CHI '06 Extended Abstracts on Human Factors in Computing Systems* (pp. 923–928). New York, NY, USA: ACM.

Judd, C., McAlliste, C. & Steel, W. (1905). General introduction to a series of studies with eye movements by means of Kinetoscopic photografs. *Psychological Review, Monograph supplements, 7*, 1-16

Just, M. A. & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*, 441-480.

Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension*. Psychological Review, 87*(4), 329-354.

Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 181–211). Hillsdale, NJ: Erlbaum.

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

Lahiri, U., Warren, Z., & Sarkar, N. (2011). Design of a Gaze-Sensitive Virtual Social Interactive System for Children With Autism. Neural Systems and Rehabilitation Engineering, *IEEE Transactions on, 19*(4), 443–452.

Land, M. (1999). Motion and vision: why animals move their eyes. *Journal of Comparative Physiology* A, 185, 341–352.

Land, M., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities. *Vision Research, 41*, 3559–3565.

Mele, M., & Federici S. (2012) Gaze and eye-tracking solutions for psychological research. *Cognitive Processing,, 13*(1), 261-165

Moore, G. (1965). Cramming more components onto integrated circuits. *Electronics, 38*(8), 114-117.

Morin, E. (1987). O método III: *o conhecimento do conhecimento*. Lisboa: Europa-América.

Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. *Emotion, 6*(2), 257–68.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General,* 130(3), 466-478.

Paivio, A. (1991). Dual Coding Theory: Retrospect and current status. Canadian *Journal of Psychology, 45*(3), 255-287.

Partala, T., & Surakka, V. (2003). Pupil Size Variation as an Indication of Affective Processing. International *Journal of Human Computer Studies, 59*(1-2), 185-198.

Porter, G., Troscianko, T., & Gilchrist, I.D. (2007). Effort during visual search and counting: insights from pupillometry. Quarterly *Journal of Experimental Psychology, 60*, 211-229.

Porterfield, W. (1737). An essay concerning the motions of our eyes. Part I. Of their external motions. *Edinburgh Medical Essays and Observations, 3*, 160-263.

Posner, M. I. (1980). Orienting of attention. Quarterly Journal of Experimental Psychology, 32, 3–25.

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

Rayner, K., & Castelhano, M.S. (2007). *Eye* movements during reading, scene perception, visual search, and while looking at print advertisements. In M. Wedel & R. Pieters (Eds.), *Visual Marketing: From attention to action.* (pp. 9-42). Lawrence Erlbaum: New Jersey, USA.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.

Rosa, P. J., Caires, C., Costa, L. , Rodelo, L., & Pinto, L. (2014) . Affective And Psychophysiological Responses To Erotic Stimuli: Does Color Matter? In P. Gamito and P.J. Rosa*, I see me, you see me: inferring cognitive and emotional processes from gazing behavior* (pp. 171-190). Newcastle upon Tyne: Cambridge Scholars Publishing

Rosa, P.J., Esteves, F. & Arriaga (2010). Attention and physiological responses to biologically relevant stimuli: An eye-tracking study using subliminal procedures. *Psychophysiology, 47*, S99

Rosa, P. J., Esteves, F. & Arriaga, P. (2012). Ver ou não ver, eis a questão. A relação entre a

emoção e a atenção visual seletiva. *In-Mind 3*(1-4), 9-18.

Rosa, P. J., Esteves, F., & Arriaga (2014). Effects of fear-relevant stimuli on attention: integrating gaze data with subliminal exposure. Proceedings of IEEE *International Symposium on Medical Measurements and Applications,* 1 , 8-14.

Rosa, P. J., Esteves, F. & Arriaga, P. (in press). Beyond traditional clinical measurements for screening fears and phobias. *IEEE Transactions on Instrumentation and Measurement*.

Rosa, P J., Gamito, P., Oliveira, J., & Morais, D. (2011). Attentional orienting to biologically fear-relevant stimuli: data from Eye Tracking using the continual alternation flicker paradigm. *Journal of Eye Tracking, Visual Cognition and Emotion, 1*(1), 22-29.

Rosa, P. J. & Morais, D. (2011). O homo emocionalis e a tomada de decisão: a irracionalidade da escolha. In F. Rodrigues, *A influência do Neuromarketing e a tomada decisão* (pp. 77-95). Viseu: Psicosoma.

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54*(2), 93–105.

Scott. W. (1908). *The Psychology of Advertising: A Simple Exposition of the Principles of Psychology in Their Relation to Successful Advertising.* Boston: Small, Maynard and Company.

Schütz, A., Braun, D., & Gegenfurtner, K. (2011). Eye movements and perception: A selective review. *Journal of Vision*, 11(9), 1-30.

Spielberger, C. (2004). *Encyclopedia of Applied Psychology (1st ed).* Oxford: Elsevier.

Steinman, R. (2003). Gaze control under natural conditions. In: Chalupa, L.M., Werner, J.S. (Eds.). *The Visual Neurosciences. MIT Press*, Cambridge MA, pp. 1339–1356.

Steinhauer, S., Boller, F., Zubin, J., & Pearlman, S. (1983). Pupillary dilation to emotional visual stimuli revisited. *Psychophysiology*, 20, 472.

Steinhauer, S., Siegle, G., Condray, R., & Pless, M. (2004). Sympathetic And parasympathetic innervation of pupillary dilation during sustained processing. *International Journal of Psychophysiology,* 52, 77–86.

Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. Quarterly *Journal of Experimental Psychology*, 59 (11), 1931 – 1949.

Thornton, T.L. & Gilden, D.L. (2007). Parallel and Serial Processes in Visual Search. *Psychological Review, 114*(1), 71-103.

Treue, S. (2001). Neural correlates of attention in primate visual cortex. *Trends in Neurosciences, 24*, 295-300.

Voßkühler, A., Nordmeier, V., Kuchinke, L., & Jacobs, A.M. (2008). OGAMA - OpenGazeAndMouseAnalyzer: Open source software designed to analyze eye and mouse movements in slideshow study designs. Behavior *Research Methods, 40*(4), 1150-1162.

Warde, A. (1999). Convenience food: space and timing, *British Food Journal, 101*(7) 518-27.

Wedel, M., & Pieters, R. (2006). Eye Tracking for visual marketing. *Foundations and Trends in Marketing, 1*(4), 231-320.

Yantis, S. (1988). On analog movements of visual attention. *Perception Psychophysics, 43*, 203–206.

Yarbus, A. (1967). Eye movements during perception of complex objects. In: L. Riggs (ed.) *Eye Movements and Vision* (pp. 171-196). New York: Plenum Press.

Young, L., & Sheena, D. (1975). Methods and designs: Survey of eye movement recording methods. *Behavior Research Methods and Instruments, 7*, 397-429.

Zhiwei, Z., & Qiang, J. (2007). Novel Eye Tracking techniques under natural head movement. *IEEE Transactions on Biomedical Enginnering, 54*(12), 2246–2260.