

## MASTER

### What behavior predicts banner effectiveness? a mouse and eye tracking study

Bremmers, L.C.M.

*Award date:*  
2011

[Link to publication](#)

#### **Disclaimer**

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

Eindhoven, November 2011

**What behavior predicts banner effectiveness?  
A mouse and eye tracking study.**

by L.C.M. Bremmers

identity number 0667307

in partial fulfilment of the requirements for the degree of

**Master of Science  
in Human Technology Interaction  
*IE&IS - University of Technology Eindhoven***

Supervisors:

Dr. ir. M.C. Willemsen (IE&IS)

Prof. dr. C.C.P. Snijders (IE&IS)

# **What behavior predicts banner effectiveness? A mouse and eye tracking study.**

## **Abstract**

The click through rate (CTR) is the number one metric of online advertising effectiveness, even though a large body of research has indicated that the CTR underestimates advertising effectiveness. These studies indicate that there are also cognitive and affective responses to banners. The current study used an eye and mouse tracker to relate these responses to their accompanying (goal-directed or exploratory) search behavior, such that it would be possible to infer responses based on search behavior and thereby construct a richer measure of online advertising effectiveness. Eighty-one participants were randomly assigned to one of four conditions (exposed or not exposed to banner ads and immediate or delayed questionnaire) and visited a regional news website for 20 minutes (10 minutes fulfilling tasks and 10 minutes of free surfing). During the website visit the participants' eye and mouse behavior was tracked. The results show that banner effectiveness goes beyond CTR. Banners received more fixations and were more often recognized than clicked upon. The most fixations and the highest recognition scores were achieved in the exploratory search mode. On the contrary to prior research, affective responses did not significantly differ between the exposed and control condition or between the exploratory and the goal-directed search mode. The search mode the user is in could be inferred by eye fixations; by looking at the number of unique AOIs visited, and by the time spent on the webpage. Mouse movements, on the other hand, could not predict the mode the user is in and neither can they predict to what area the user is paying attention to. Therefore a mouse tracker alone is insufficient to infer cognitive and affective responses. Making it impossible for now to record these responses unobtrusively in practice.

**Keywords: websites, advertising, attention, eye movements, human information processing, cognition, memory**

## **1. Introduction**

The internet is by far the most important medium in the lives of European consumers (Digital influence index study, 2008 in Klever, 2009). Moreover, internet advertising is a billion dollar industry at which the amounts are growing at incredible rates (e.g. Jaworska & Sydow, 2008). Nevertheless, internet advertising still needs a lot of research to understand how online advertising works and how performance can be measured.

At the moment the click through rate (CTR) is the number one metric for online advertising effectiveness, even though a large body of research has indicated that the CTR underestimates advertising effectiveness.

Briggs and Hollis (1997), and Drèze and Hussherr (2003) show empirically that banner advertising has a significant impact on consumers independent of click-through.

Drèze & Hussherr (2003) describe that traditional advertising benchmarks (i.e. brand awareness, unaided advertising recall, brand recognition and aided advertising recall (Tellis, 1998)) have proven to be effective for online advertising as well and even work better than CTR.

Briggs and Hollis (1997) results show that the CTR does not reflect the overall value of banner advertising. In their study they have used a system called the “BrandDynamics” system, which consists of two parts; The Consumer Value Model, which identifies the probability that an individual will choose a particular brand for their next purchase (i.e. Consumer Loyalty), and the BrandDynamics Pyramid, which explains the variation in respondents' Consumer Loyalty scores. The higher the level of the pyramid, the higher the attitudinal involvement; from knowing that a brand exists to finding it difficult to consider another brand. Briggs and Hollis found large differences in the scores of three different banners, while each banner received approximately the same CTR. This indicates that the CTR does not reflect the overall value of a banner exposure, and therefore the use of CTR alone undervalues the web as an advertising medium.

Also a study by the Internet Advertising Bureau (1997, in Lapa, 2007) on over 16,000 users found that online advertising can generate increases in advertisement awareness, brand awareness, product attribute communication and purchase intent. Moreover, nearly all of the impact was generated without a click-through to the advertiser’s website.

All these studies show that it is recommended to use other measures next to CTR in order to measure effectiveness more reliably. However, because these measures cannot be measured unobtrusively, this information is not used in practice. Yet research indicates that some responses appear more often when individuals are in a goal-directed search mode and others appear more often when individuals are in an exploratory search mode. So it might be possible to predict the kind of response based on the mode the individual is in.

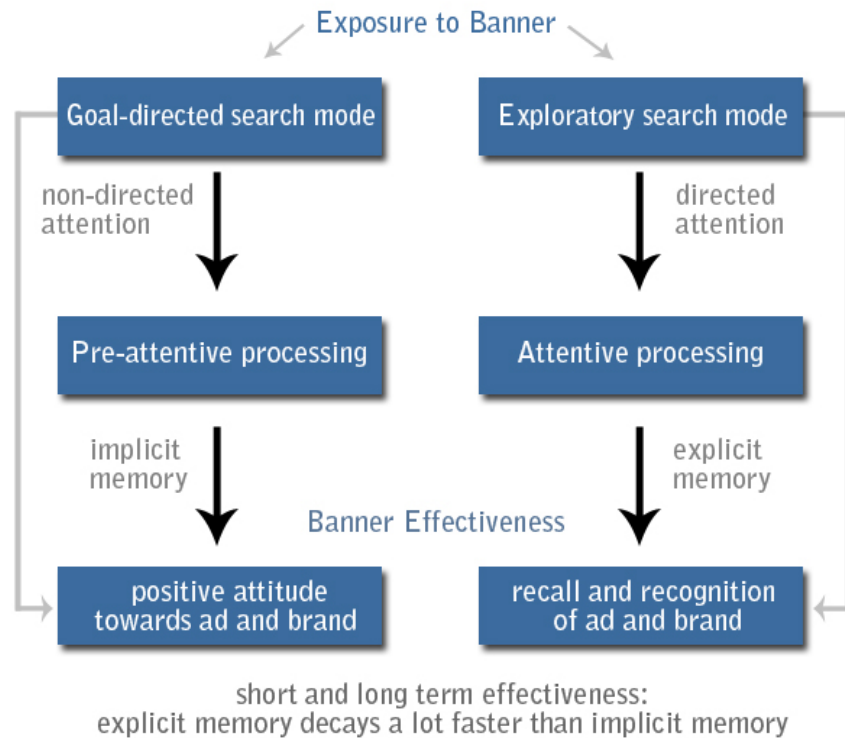
There are two kinds of search modes; the goal-directed search mode (i.e. searching for specific information) and the exploratory search mode (i.e. exploring webpages without any direct, pre-determined goal) (see figure 1.1).

Depending on the mode individuals are in, the way banner ads are processed differs. Users in a goal-directed search mode spend a lot of attention on their task and, because human’s attention resource is limited (Wang & Day, 2007), little to no attention is left for peripheral stimuli; in this case banners. Banners are irrelevant to the search task at hand and tend to slow down the process and are therefore (sub-consciously) avoided (Chatterjee, 2008). This phenomenon is also known as “banner

*“The CTR value [...] denotes the ratio between the number of times the ad was clicked by a user and the number of times the ad was presented on a web page (impressions). For example, CTR of 3% means that the ad was clicked in 3% of the cases it was presented to the web users.”*

(Drèze & Hussherr, 2003)

blindness” (e.g. Benway & Lane, 1998; Drèze & Hussherr, 2003; Albert, 2002). Because of this banners have a higher likelihood to be processed pre-attentively (i.e. users are not consciously aware of the banner ads and process them automatically without paying any attention to them) in this mode (Yoo, 2007).



**Figure 1.1** The two search modes, their accompanying way of processing and responses

Individuals who are in an exploratory search mode spend less attention to the content and peripheral stimuli can compete for attention, as attention is not so singularly focused (Janiszewski, 1998). Because of this the banners have a higher likelihood to be processed attentively in this mode. Consistent with this, Lapa (2007) found higher dwells and total viewing time of banners in the exploratory search condition as compared to the goal-directed search condition.

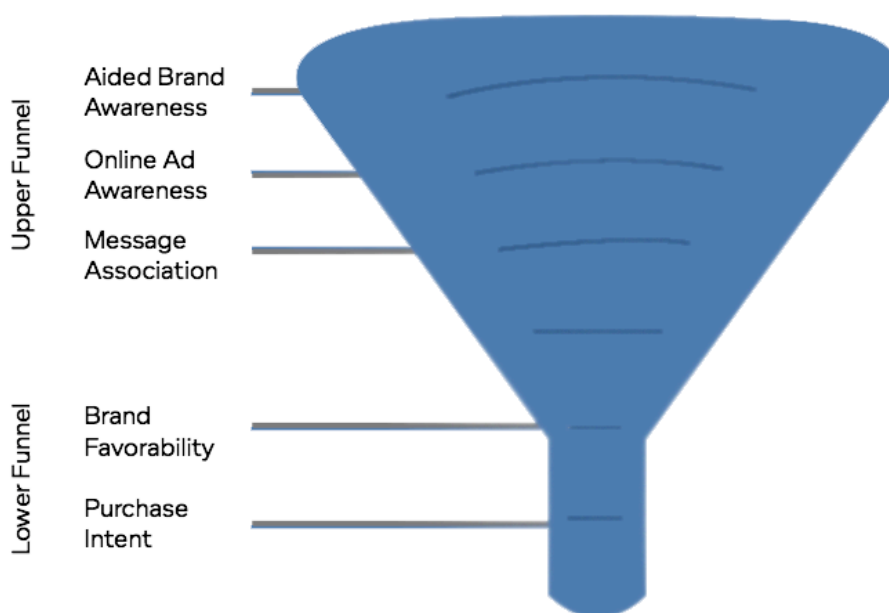
The way banners are processed affects the effectiveness of the banner. Previous research shows that both pre-attentively and attentively processed banners have positive effects, but the effects are different; a study by Goodrich (2011) indicates that the amount of attention spend on an ad correlates with memory (i.e. recall and recognition). In line with this, Danaher and Mullarkey (2003) and Pagendam and Schaumburg (2001 in Calisir & Karaali, 2008) found that web users in an exploratory search mode are more likely to recall and recognize banners than subjects who search for information in a goal directed way.

On the other hand, banner ads which are processed at the pre-attentive processing level influence the users’ liking of the ad and the brand in a positive way (Janiszewski, 1990a, 1990b, 1993 in Drèze & Hussherr, 2003; Vanhuele et. al, 2005; Yoo, 2008; Kunst-Wilson & Zajonc, 1980;

Lewicki, Hill, & Czyzewska, 1992, in Yoo, 2008). Thus nonetheless pre-attentive processed ads receive limited attention, they can still effectively change attitudes (Anand, Holbrook, and Stephens 1988; Janiszewski 1988, 1990a, in Janiszewski, 1993). Moreover, attention was found to be negatively related to attitudes (e.g. Bornstein, 1989 in Goodrich, 2011; Heath, Brandt & Nairn, 2006; Goodrich, 2011), this explains why a lower level of attention is found to be more effective for brand building (Yoo, 2005 in Goodrich, 2011; Heath, Brandt & Nairn, 2006 in Goodrich, 2011; Goodrich, 2007 in Goodrich, 2011). Thus an attentive processed ad should receive a lower brand attitude compared to an ad that is pre-attentively processed or not processed at all (i.e. the ad is not shown) and a pre-attentive processed ad should receive a higher brand attitude compared to an ad that is attentively processed or not processed at all.

The reason why pre-attentive processed ads influence the user’s liking of the ad and brand in a positive way is that mere exposures generate a feeling of familiarity that is later interpreted as a preference for the stimulus (Mandler, Nakamura, and van Zandt, 1987; Zajonc, 1980 in Janiszewski, 1993). This might explain why individuals who have processed an ad pre-attentively have a greater likelihood to select the advertised brand in a choice situation and to include the brand in a consideration set than individuals who are not exposed to the ad (Yoo, 2008).

Which way of processing is more effective for banner ads is unclear. First a definition of what banner effectiveness refers to in this experiment; When we look at the goals of advertisers we see that ultimately, advertisers are interested in sales (Hoffman & Novak, 2000). To get to sales, first the initial stage of brand awareness has to be completed to continue to message association, brand favorability, purchase intention, and, ultimately, purchase (DoubleClick, 2009 report). Therefore the higher the stage in this purchase funnel (see figure 1.2) that is positively affected by the banner, and the greater this effect is, the better the banner effectiveness.



**Figure 1.2** Purchase funnel

The first part of this study is designed to find out which kind of processing (pre-attentive versus attentive) is better in terms of banner effectiveness, both on the long and short term, as prior research indicates that delay inhibits explicit memories (which are created by attentive processing) but does not inhibit implicit memories (which are created by pre-attentive processing) to a great extent (Richardson-Klavehn & Bjork 1988; Schacter, 1987). For example, prior research has shown that both recall and recognition (which are explicit memory tests) are negatively affected by time delay (e.g. Singh, Rothschild, and Churchill, 1988; Chatterjee, 2008), with recall lower than recognition since you do not have to recall the ad for recognition to occur (Chattopadhyay & Nedungadi, 1992 in Chatterjee, 2008). Furthermore, Tulving, Schacter and Stark (1982) and Vanhuele et. al (2005) found that implicit memories are preserved even after a week's delay. Also DeSchepper and Treisman (1996 in Yoo, 2008) found that implicit memory is quite durable over time, even if there was only a single exposure to a stimulus.

Once it is known what kind of processing is more effective for banners, it would be great to use this information in practice as well. If it is possible to infer in which mode a user is, the kind of processing can be inferred and thereby banner effectiveness could be predicted. Therefore the second part of this study is concerned with how to distinguish between the two different modes (i.e. goal-directed search mode and exploratory search mode). This study tries to infer the modes by using eye and mouse tracking data. An eye movement is highly related to attention (Just & Carpenter, 1976; Kahneman, 1973, in Wang & Day, 2007), and is considered to represent attention itself. Eye tracking will provide us with rich data of where the user's attention is directed to (Theuner, Pischke & Bley, 2008; Just & Carpenter, 1976 in Poole & Ball, 2005; Wedel & Pieters, 2007). For example the spatial density of fixations can indicate whether someone is focused and searches efficiently (Cowen, Ball & Delin, 2002 in Poole & Ball, 2005). Further the total number of fixations is thought to be negatively correlated with search efficiency (Goldberg & Kotval, 1999). Also the saccade/fixation ratio can tell us whether the user spent more time searching or more time processing (Goldberg & Kotval, 1999). It is expected that users in a goal-directed search mode will search more focused and efficient and spend more time processing and less time searching than users in an exploratory search mode.

Next to eye tracking, also mouse tracking is used. Prior research has found a strong correlation between eye movements and mouse movements (Chen, Anderson & Sohn, 2001; Egner, Itti & Scheier, 2000; Cooke, 2006; Drunen, Broek & Heffelaar, 2009). Chen, Anderson and Sohn (2001) even argue that a mouse device would be a very good alternative to an eye tracker. This implies that the mode in which user is might be inferred by mouse data only, and thereby the effectiveness of the banner could be predicted in practice as well because -as opposed to eye movements- mouse movements can be collected accurately, easily, remotely, and on a large scale, using Javascript (Rodden, Fu, Aula & Spiro, 2008).

To summarize; the goal of this study is twofold; first this study tries to find out what effect behavior (i.e. goal-directed search versus exploratory search) has on banner effectiveness via the corresponding

kind of processing (i.e. pre-attentive or attentive). Second this study tries to find out if this behavior can be inferred with eye and mouse tracking and thereby, in extension, predict banner effectiveness.

The main contribution of this study compared to prior research in this field is that the ecological validity is better. This study tries to keep the web browsing experience as realistic as possible by using an existing website where the user can browse each page as long (to some extent) as she wants without restrictions, whereas other studies; constrained the user or the web browsing experience; for example Goodrich (2011) made participants click on whatever caught their eye on a series of eight static webpages, each presented for 10 seconds. Another example is the study of Yoo (2008) in which he showed three webpages for 45 seconds each and gave participants the task to either evaluate the webpages (i.e. directed attention to banners) or to read the contents of the webpages (i.e. non-directed attention to banners). Also Lapa (2007) and Danaher and Mullarkey (2003) showed webpages for a fixed amount of time (i.e. 15 seconds in Lapa's study and 20, 40 or 60 seconds in Danaher and Mullarkey's study). Besides Lapa and Danaher and Mullarkey did not allow participants to click or navigate to other pages.

Another study with low ecological validity is the one of Vanhuele, Courbet, Lavigne and Borde (2005), who placed the banner on the webpage 5 seconds after participants had started reading and again removed the banner from the page after three seconds. Further most studies made up mock websites and/or content (e.g. Janiszewski, 1993; Danaher & Mullarkey, 2003; Shapiro, 1999). Next to a more realistic browsing experience, also the banners in this study are as realistic as possible and contain existing foreign brands that most people in the Netherlands are unfamiliar with whereas other studies use fictitious brands and ads. Also a diverse set of banners is used to overcome participants motivations for certain product categories (e.g. Goodrich (2011) only used one advertisement of an electric shaver).

An eye tracker is used to measure pre-attentive and attentive processing in an objective way, whereas most prior research used subjective measurements as introspection and self-report. For example, Drèze and Hussherr (2003) used questions about recall, recognition and awareness of banner advertising to indicate if advertisements were processed attentively or pre-attentively. Further Janiszewski (1993), Shapiro and MacInnis (1992) and Yoo (2008) assumed that banners in the peripheral area were processed pre-attentively when participants were given a task to fulfill (e.g. reading a mock newspaper in the case of Janiszewski). They also verified this by post-hoc checking the recognition levels for the banners.

This subjective way of measuring attention to banners is prone to error because human's recall of banners can fail (e.g. Biel, 1993, in Wedel & Pieters, 2007). Nonetheless recall is a frequently used and an accepted way to measure attention (Wedel & Pieters, 2007). However, Wedel and Pieters also mention that it is not the best way; it would be better to measure attention in an objective way as well.

Vanhuele et al. (2005) made use of eye tracking in order to make sure banners were not consciously processed, by removing banners from the page as soon as the eye came to close to them. However this study does not reflect a realistic web browsing experience and does not look into banners which are consciously processed.



Another study that made use of an eye tracker to trace where participants were looking at and to find differences in attention to banners between search modes is that of Lapa (2007). The problem with his study is that the banners he used had a different format and place than they usually have; the banners were not placed in the periphery, the banners were all for in site promotions, none of them included a brand name and most important; participants in the task-free part were instructed to “look at the website” and could not browse or click whereas participants could browse and click in the task part. Therefore his conclusions cannot be generalized.

Moreover, this is the first experiment that tries to extract the modes the user is in by eye fixations and mouse behavior, and that tries to connect behavior to banner effectiveness and thereby construct a richer measure of online advertising effectiveness. Based on such a measure, a fairer payment system could be developed that is not solely based on CTR. Currently, a lot of publishers have a disadvantage because they only get paid for the number of clicks while banners are also effective without clicks.

The current experiment includes a number of tasks (which represent the goal-directed search mode) and free surfing tasks (which represent the exploratory search mode) which participants have to fulfill while visiting a regional news website. Immediately after the experiment, or after a week’s delay, participants get a questionnaire with questions about the banners and the advertised brands. During the experiment mouse and eye behavior is tracked.

It is expected that banners have a greater likelihood to be processed attentively in the exploratory search mode and have a greater likelihood to be processed pre-attentively in the goal-directed search mode; banners which are processed attentively are better recognized, but this effect decays over time and banners that are processed pre-attentively receive better affective responses and this effect lasts (longer) over time.

Finally it is expected that there are differences in behavior between the two search modes; behavior is expected to be more efficient and focused in the goal-directed search mode. Hereby the mode the user is in could be inferred based on the shown behavior. Also it is expected that eye fixations and mouse movements are highly correlated and therefore a mouse tracker would also be able to tell where the user is paying attention to. It is expected that mouse behavior alone should be sufficient to distinguish between modes.

## 2. METHOD

### 2.1 Design

This study used a mixed design (see table 2.1) with two between-subjects independent variables; exposure (with two levels: exposed to experimental banner ads or not exposed to experimental banner ads) and time of questionnaire (with two levels: immediately after the experiment or after a week's delay (consistent with Shapiro & Krishnan, 2001; Chatterjee, 2008; Chattopadhyay & Nedungadi 1992 in Chatterjee, 2008)). There was one within-subjects independent variable; task (within-subjects, with two levels: free surfing or search task) and five dependent variables (each measured ten times; one time for each banner); brand attitude, brand in consideration set, purchase intent, brand favorability and ad awareness (i.e. recognition).

**Table 2.1** Mixed Design: independent variables

<b>Between-subjects</b>	Exposed to banners	<i>OR</i>	Not exposed to banners
	Immediate questionnaire	<i>OR</i>	Delayed questionnaire
<b>Within-subjects</b>	Tasks	<i>AND</i>	Free surfing tasks

### 2.2 Participants

There were eighty-one participants in total of which 28 in the exposed to banners, immediate questionnaire group; 26 in the exposed to banners, delayed questionnaire group; 15 in the control (not exposed to banners), immediate questionnaire group and 12 in the control, delayed questionnaire group. Most of the participants were students of the University of Technology in Eindhoven, the Netherlands. About half of the participants was male and half was female; 44 males and 37 females. The age of the participants ranged from 18 to 33 years ( $M=22,3$ ,  $SD=2,8$ ) and their self-reported internet usage ranged from 1 to 9 hours a day ( $M=4,1$ ,  $SD=2,0$ ).

Participants signed up for participation in the experiment voluntarily and were rewarded with €7,50 if they were from the University of Technology in Eindhoven or €9,50 if they were not. Participants were not told about the purpose of the study, but were made to believe it was a usability study for the website involved. Seventy-three participants were unfamiliar with the involved website, five participants had heard of it and two participants visit the website occasionally.

## 2.3 Apparatus and Materials

Participants visited the website using Internet Explorer 9 (without any limitations) on a desktop computer (OS: Windows 7) with a resolution of 1600 x 1200 pixels.

A Tobii X120 table mounted eye tracker was used to track participants eye movements and fixations during the website visit. To track participants mouse movements during this visit, the Mouse Recorder Pro 2 by Nemex Studios was used.

## 2.4 Procedure

This experiment was conducted in a controlled lab setting at the University of Technology in Eindhoven.

Before starting with the experiment, the eye tracker was calibrated to the participant's eyes. After calibration tasks were assigned to the participants, one task at a time. All these tasks had to be fulfilled by visiting a regional news website; [www.kliknieuws.nl](http://www.kliknieuws.nl). The tasks consisted of 5 search tasks and 2 free surfing tasks. The order of the search tasks was randomly chosen, the free surfing tasks were fixed after 2 and again after 3 search tasks (see figure 2.1 for an overview of the procedure).

Each search task took a maximum of 2 minutes (10 minutes for all search tasks together) and each free surfing task took 5 minutes (10 minutes for the free surfing tasks together).

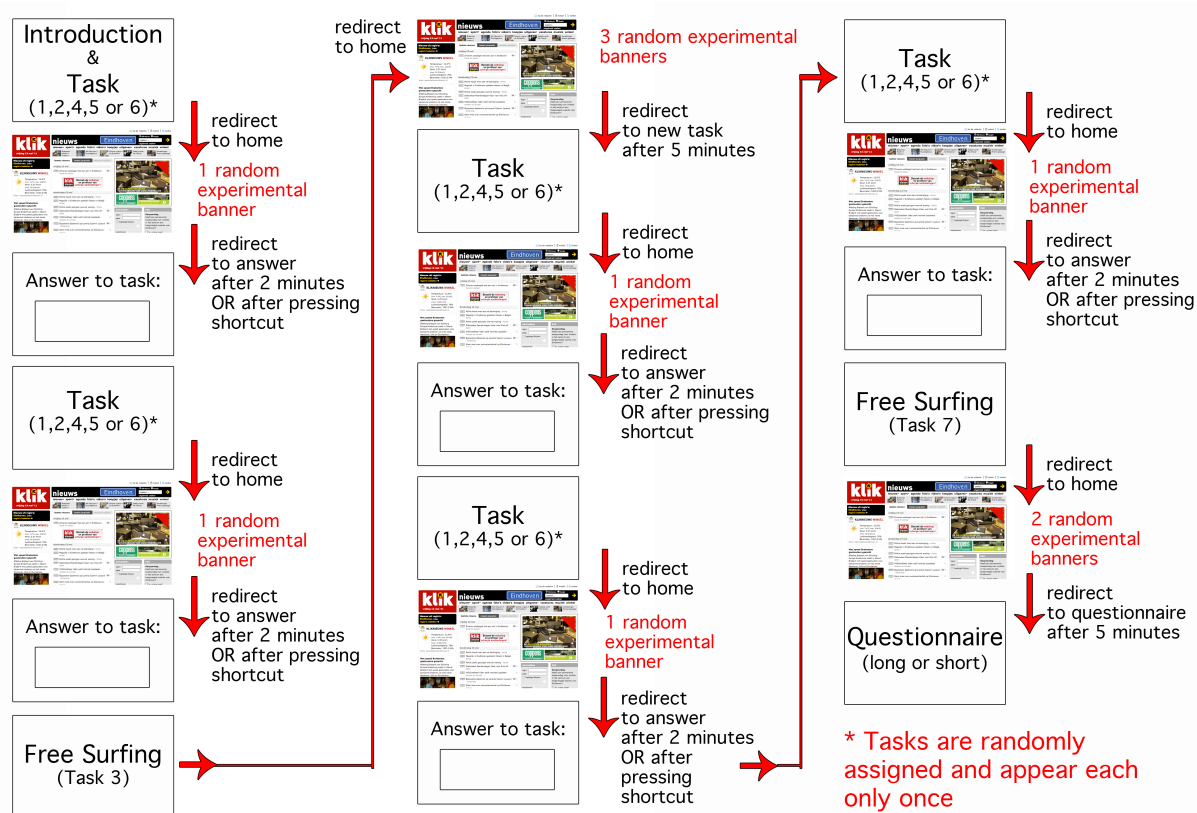


Figure 2.1 Experiment procedure

### **2.4.1 Tasks**

Participants were given 2 minutes to complete a search task. This time was believed to be a reasonable amount of time to find the answer; not too short to frustrate the user and not too long to be bothersome. After the 2 minutes a full screen message popped up where subjects could either fill out the answer to the question or they could indicate that they had not found the answer. When participants found the answer before the 2 minutes had elapsed, they could press a shortcut (right arrow on the keyboard) to immediately go to the answer field. This shortcut was included to make sure participants would not switch from the goal-directed search mode to the exploratory search mode when they had found the answer, else this would obstruct the task manipulation.

After the answer had been filled out, a new task was shown on the screen (In the case of the previous task being a free surfing task subjects were only served with the new task).

The search tasks users got were information search tasks. An example of a search task used in this experiment is: “you want to do something fun next week. Search for an event that interests you and specify the date, time and place where the event takes place.” All tasks represented realistic search tasks for information that users could be looking for in real life as well. This because participants who are given specific tasks to perform, that accurately reflect a realistic situation, will behave differently from participants who are given atypical tasks to perform that they would not otherwise perform outside the laboratory environment (John & Marks, 1997, in Lapa, 2007).

### **2.4.2 Banners**

In this study ten different experimental banners (see appendix 1) were used, some of them matched student interests and some of them did not. For participants in the exposed condition, there was one experimental banner in each search task and two or three experimental banners in the free surfing tasks. Each banner was shown only once and each banner was shown on a different web page. The order of the impressions was randomly chosen, but each advertisement was shown, and each ad was shown only once. For the remaining pages filler ads were shown (see appendix 2).

Because Wang and Day (2007) found that attention for banners is highest at the beginning and at the end of what they call a ‘meaningful path’ (e.g. reading a news story that is divided over a number of pages) the experimental banners were placed on different pages in time to control for this effect. This could not be done completely randomly because the experimental ad has to be somewhere within the minimum number of pages that is usually visited, to make sure it is shown before the task ends. Therefore the experimental banners were shown at page 1, 2 or 3 in the case of a search task and at page 1,2,3,4 or 5 in case of a free surfing task (more pages because of more time). The pages that did not include an experimental banner had a filler ad, of the same size and on the same place, instead. Participants in the control condition were only exposed to filler ads.

Banners and banner placement have been deliberately chosen. Research has shown that different advertisements can have different effects on users; amongst others the size of the banner, the product or service advertised, the design of the banner and the content have effects on users;

Size and place have an effect on the attention banners receive; the bigger the ad the better (Wedel & Pieters, 2007; Janiszewski, 1998) and banners which are located at the top of a page receive more attention than banners at the bottom (Rosenkrans, 2009). Therefore the banners used in this experiment all have the same size and were all displayed at the same place on the web page.

Chtourou, Chandon and Zollinger (2002) found that mentioning the price of a product or service in an advertisement reduced direct response. In line with this, Rettie, Grandcolas and McNeil (2004) discovered that banners which mentioned neither price nor promotional offers had the most effect on click-through and post-impression rates. Price nor promotional offers are mentioned in the used advertisements.

Previous research shows that there is an effect of animation in banners; animated banners are more difficult to remember than static lookalikes (Burke, Hornof, Nilsen & Gorman, 2005; Burke, Gorman, Nilsen & Hornof, 2004). The Doubleclick 2009 report remarks that static banners work better than their animated counterparts. Therefore only static banners are used in this study.

Wedel and Pieters (2007) have found that participants fixated color before non-color ads, noticed more color ads than non-color ads, and viewed color ads 21% longer than equivalent ads without color. In this study color is used in all ads.

The brands which are depicted on the banners are existing foreign brands. Using existing brands is not advised; previous research (Dahlen, 2001, in Chatterjee, 2008; Yoo, 2008; Vanheule, et al., 2005) has indicated that prior familiarity with the advertised brands could potentially confound the results (consistent with Danaher & Mullarkey, 2003; Yoo, 2008; Vanhuele et. al 2005; Wedel & Pieters, 2007), but since this study has used foreign brands unfamiliar to (most) Dutch participants this can be considered similar to using non-existing brands. The advantage of using existing brands is that more realistic brand names, logos and ads can be used.

### **2.4.3 Website**

Eindhoven, the area in which the University of Technology is situated, is covered by the Kliknieuws website. Besides Eindhoven, other cities and villages in the same province (Noord-Brabant) are covered. Participants could select of which region they would like to read the news. The website contents were exactly the same as those of the original website, at the moment of the experiment, with the only exception of the manipulated banner ads. The participants could access the entire website without restrictions.

The Kliknieuws website has several spots for advertisements. In this study the focal advertisement spot (see figure 2.2) was used to display the experimental and filler ads. The banners have a size of 336 x 280 pixels. All other (smaller) banners were kept the same as on the original website at the moment of the experiment.



Figure 2.2 Screenshot from the Kliknieuws website with indication of the place for banners

## 2.5 Measurements

During the experiment several measures were recorded by the eye and mouse tracker. Afterwards a questionnaire recorded the remaining measures.

### 2.5.1 Eye- and mouse tracker

Various measures were recorded during the experiment. Per page the following was recorded:

- The task the participant was performing.
- The ad that was active.
- The mapped fixation point (in X and Y coordinates).

- The number of fixations participants made per Area of Interest (AOI); more fixations on a particular area indicate that it is more noticeable, or more important, to the viewer than other areas (Poole, Ball & Phillips, 2004 in Poole & Ball, 2005; Russo, 2010; Theuner, Pischke & Bley, 2008).
- The amount of time each fixation lasted; longer fixation durations indicate difficulty in extracting information, or it means that the object is more engaging in some way (Just & Carpenter, 1976 in Poole & Ball, 2005; Goldberg & Kotval, 1999; Russo, 2010).

Wang and Day (2007) indicate that either the length of time a subject's eyes are fixated on an object or the number of times she fixates on that object can be used as a measure of attention because these indicators are highly correlated. In this study the number of times a participant fixates on an AOI is used.

- The time of each page visit; Danaher and Mullarkey (2003) indicate that the longer a person is exposed to a web page containing a banner advertisement, the more likely they are to remember that banner advertisement.
- The location of the mouse (in X and Y coordinates).
- The delay in time between the current and the previous mouse position.
- Keyboard input.
- Mouse clicks and scrolling.

## 2.5.2 Questionnaire

The questionnaire recorded implicit and explicit memory measures, the “purchase funnel responses” (e.g. brand attitude, brand favorability etc.) and some control questions.

### *Implicit memory measures*

To subjectively measure whether a banner has been processed attentively or pre-attentively cognitive measures have been employed. To measure implicit memory the process dissociation procedure (PDP) was employed (consistent with Shapiro & Krishnan, 2001; Yoo, 2007; Yoo, 2008). PDP is used to parse out the effects due to conscious memory retrieval, yielding an unbiased estimate of the amount of automatic influence caused by implicit memory retrieval. PDP consists of two tasks; an inclusion task to measure automatic influences and an exclusion task to measure both conscious and automatic influences. Both tasks are word stem completion tasks with words that correspond to the advertised products in the banners. For example for a banner which advertises glasses, the corresponding exclusion task could be; complete the word stem gla\_\_ with words that do not correspond with the shown banners. This way participants who processed the banner unconsciously have an increased likelihood of completing the word stem with “glasses” because of automatic influences and participants who processed the banner consciously have a decreased likelihood of completing the word stem with “glasses”. So the exclusion task taps into automatic influences.

Because automatic influences are overwritten by conscious memory retrieval there is also the inclusion task. Contrary to the exclusion task, the inclusion task asks for completing the word stem with words that do correspond with the shown banners. Hereby there is an increased likelihood of

completing the word stem with the right word if conscious memory retrieval is successful or if conscious memory retrieval failed but automatic influences created a correct response.

After both tasks have been completed conscious influences and automatic influences can be calculated by using the following formulas;

$$C = \text{inclusion performance} - \text{exclusion performance}$$

$$A = \text{exclusion performance} / (1 - C)$$

$$\text{Inclusion} = C + (1 - C) A$$

$$\text{Exclusion} = (1 - C) A$$

In this experiment 14 words were selected and were equally divided over the two tasks. The words that were chosen for the experiment had first been pre-tested to make sure the baseline completion rates would not be too low or too high. Selection was based on the criteria used in prior studies (e.g., Tulving, Schacter, & Stark, 1982; Yoo, 2008); words with no higher than 46% completion rates and no lower than 15% completion rates were selected.

### ***Explicit memory measures***

To measure explicit memory a recognition task was used. Studies often use recognition and recall as explicit memory measures, however recognition is a more robust measure than recall and it also measures the presence of a memory trace for the banner (Ewing, Napoli & Du Plesses, 1999 in Calisir & Karaali, 2008). Shapiro and Krishnan (2001) found that recall scores were too poor after a week's delay. Therefore this study has used recognition only. The recognition task consisted of a two-alternative forced choice task (consistent with Chatterjee, 2008; Kunst-Wilson & Zajonc, 1980 in Yoo, 2008; Shapiro & Krishnan, 2001); participants were given two alternative advertisements of which they had to make a forced choice ("Click on the banner that you recognize having seen"). The banners which were used have been pre-tested first and the banner combinations that were the most difficult (i.e. in which the correct alternative was chosen the least times) have been chosen for the subsequent experiment.

Besides subjects were asked how confident they are from 50% to 100%<sup>1</sup> that their answer is correct (consistent with Calisir & Karaali, 2008). This way the data can be transformed taking into account the confidence level of the participant (Young, 1990 in Calisir & Karaali, 2008). The recognition scores range from 1 to 12 after transformation (i.e. 1 for 100% certain and picking the wrong banner, and 12 for 100% certain and picking the right banner).

---

<sup>1</sup> participants have a 50% probability of picking the right banner to begin with.



### ***Affective & behavioral measures (Purchase Funnel)***

Brand awareness was tapped with the question: “Have you heard of [brand] of [product category]?” With answering options: Have heard of / Seen; Have not heard of / Seen; Not sure (consistent with the Doubleclick, 2009 report).

Brand attitude was measured by: “I consider [brand] to be: ...” with three 9-point semantic differential scales anchored by “positive/negative,” “good/bad,” and “favorable / unfavorable” (consistent with MacKenzie, Lutz, & Belch, 1986). Brand attitudes are expected to be highest for banners which are processed at the pre-attentive processing level, lowest for banners which are processed at the attentive processing level and intermediate for banners that are not processed at all (i.e. control condition, in which the banner has not been shown).

Purchase intent was captured by the question: “How likely is it that you would purchase [brand] when you want to buy a [specific product]?” Scale from very unlikely to very likely. (consistent with Haley & Case, 1979; Doubleclick 2009 report; Goodrich, 2011). Purchase intent is expected to be greater for banners which are processed at the pre-attentive processing level compared to conscious processing or no processing at all (in the case the banner has not been shown to the participant).

Whether an item is within the consideration set is measured by asking which brands participants would consider when they would be looking for a product of the specific product category. A list of 6 foreign brand names (which are unfamiliar to most Dutch individuals), amongst which the advertised brand name, was given to participants. The order of this list was random to control for order effects. Inclusion of a brand name in a consideration set is a necessary condition for choice (Yoo, 2008). Yoo (2008) and Shapiro (1999) found that the likelihood that an advertised product would be included in the consideration set is greater in the pre-attentive processing condition than in the control group.

Brand favorability is measured by letting participants make a top 3 of the advertised brand and two competitors. The order of the brand names was random. Brands are expected to have a better rank in the top 3 for banners which are processed at the pre-attentive processing level compared to conscious processing or no processing at all (in the case the banner has not been shown to the participant).

### ***Control questions***

To control for confounding variables some control questions were included. These are about the familiarity with [www.kliknieuws.nl](http://www.kliknieuws.nl) and the advertised brands, gender, age and internet use.

The complete questionnaire can be found in appendix 3.

### 3. Results

First it is analyzed what effect the two search modes have on banner effectiveness in terms of the purchase funnel. Second it is analyzed whether these search modes can be inferred by eye and mouse behavior.

To easily analyze the different parts of the page the participants are looking at, the screen was divided into Areas Of Interest (AOIs) (e.g. Theuner, Pischke & Bley, 2008) (see figure 3.1). The AOI of special interest is the area in which the banner is located. The AOIs are based on the six functions of Nielsen & Tahir (2002); Welcome and site identity (i.e. the Kliknieuws logo), Navigation (i.e. the five different menu bars), Content (i.e. left, middle and right content areas), Advertising (i.e. the focal banner) and sponsorship (not applicable), Self-promotion (not applicable), and Filler (i.e. the bar with “nieuws” and “Eindhoven” in it and the white background with no content).



**Figure 3.1** Areas Of Interest, each rectangle representing a different AOI.

Six incomplete data sets (of which five due to eye tracker crashes, and one due to an unfinished questionnaire) were excluded from analysis.

All scores are rounded off to two decimal places throughout and all p-values are indicated with \* for  $p < .05$ ; \*\* for  $p < .01$  and \*\*\* for  $p < .001$ .

### 3.1 Purchase funnel

#### 3.1.1 Recognition is significantly higher for the exposed, immediate questionnaire condition, the free surfing condition and for banners placed at the middle of a sequence of pages

Recognition scores are expected to be higher for the exposed versus control condition (where no banners were shown) and higher when assessed directly rather than with a delayed questionnaire. Recognition scores are expected to be higher for banners which are shown during the free surfing condition compared to the task condition. Furthermore recognition scores are expected to be higher at the beginning and at the end of a task. Therefore it is expected that the recognition scores follow a U-shaped curve over the pages visited.

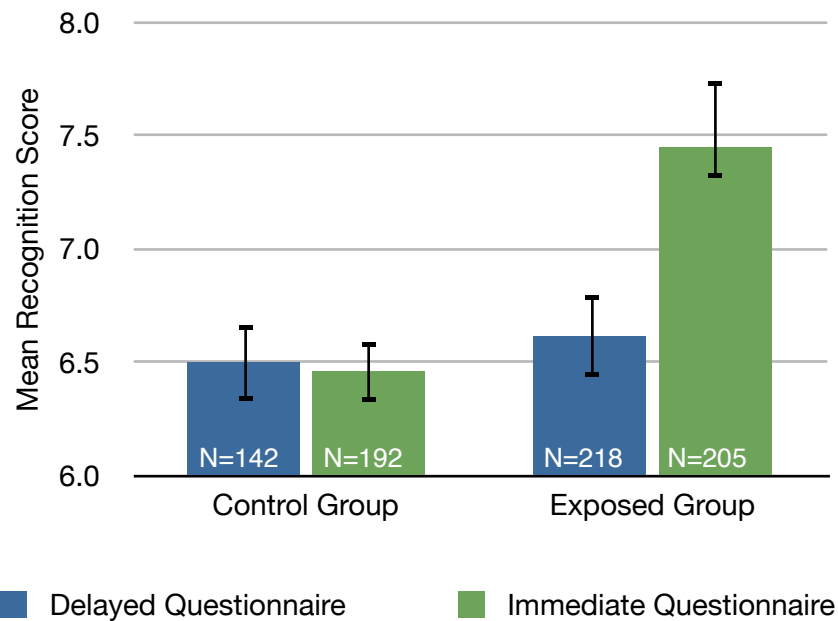
Recognition was captured by a combination of two measures; choosing the shown (i.e. correct) banner or the not shown (i.e. wrong) alternative and the degree of certainty (from 50% to 100%) of picking the correct banner. Recognition scores ranged from 1 (100% certain and wrong answer) to 12 (100% certain and right answer), with a baseline guessing score of 6,5 (see table 3.1 for an overview of the scores).

**Table 3.1** Recognition scores range

1	2	3	4	5	6	7	8	9	10	11	12
Wrong 100%					Wrong 50%	Correct 50%					Correct 100%

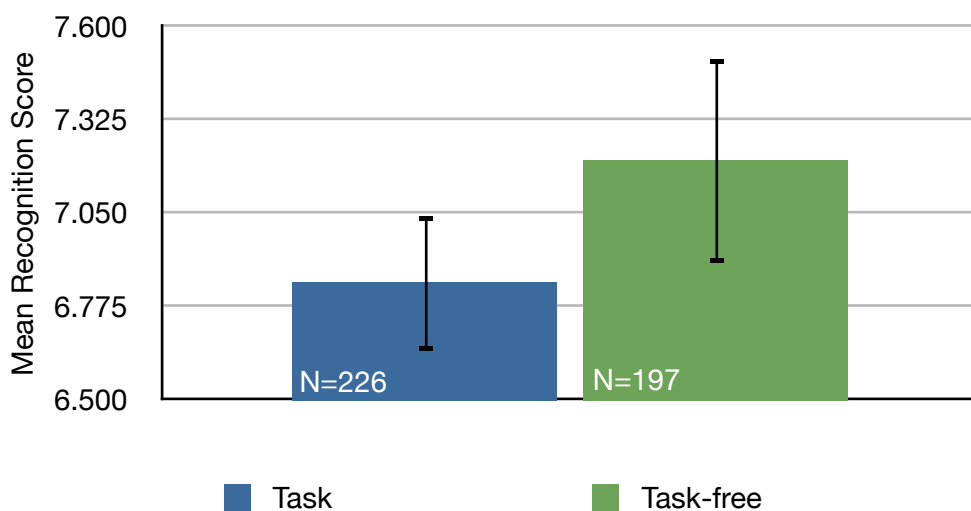
There were 10 measurements of recognition per participant (each measurement concerned one of ten banners).

The conditions to which participants were assigned to, were recalculated for each banner after the experiment; when a participant was in the exposed condition and a banner did not show up due to an error, her condition was changed -for that single banner- to control condition. This affected less than 10% of the cases.



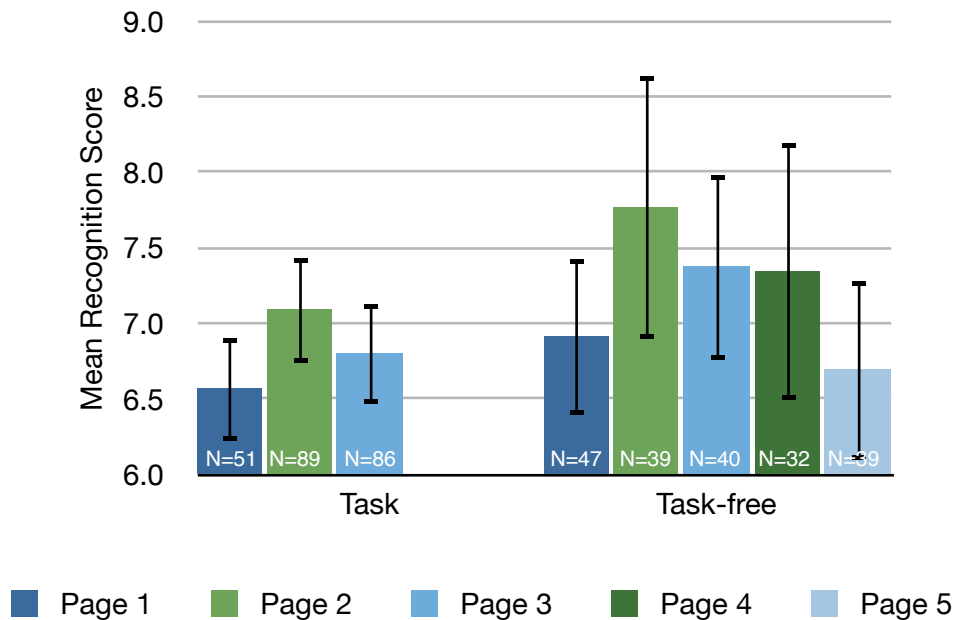
**Figure 3.2** Mean recognition scores for exposed vs control and immediate vs delayed questionnaire

The recognition scores for the exposed versus control condition and the immediate versus delayed questionnaire condition (figure 3.2) show that there is only one condition that has recognition scores that are significantly higher than the baseline level; the exposed, immediate questionnaire condition. This condition has a mean score that is significantly higher ( $p < 0.001$ ) than all other conditions. The mean recognition score of the control group is around 6.5 (the expected baseline guessing score) for both the delayed and immediate questionnaire condition. This indicates that there was no (preferential) bias towards either the used banners or their alternatives, which supports that the recognition task was well designed. The exposed group with the delayed questionnaire has a mean above the baseline level of 6.5, however it is not performing significantly better than the control group ( $p < 0.16$ ).



**Figure 3.3** Mean recognition scores for task vs free surfing condition (exposed condition only)

Looking at the free surfing and task conditions (figure 3.3) we see that as expected, the mean recognition score of the task-free condition is higher than that of the task condition. A t-test confirmed that this difference is significant ( $t(421)=-2.01, p<0.05$ ).



**Figure 3.4** Mean recognition scores by page number and task vs free surfing (exposed condition only)

When looking at the differences in recognition scores by page number (figure 3.4) we see that the scores at the middle are higher than the ones at the first and last pages. Further the deviations from the mean are a lot larger for the free surfing condition as they are in the task condition, which can be explained by the fact that participants were free to surf the website and could be doing several things, whereas participants in the task condition were all searching for the answers to the questions.

To explore the relative contribution of the factors discussed they are regressed on the recognition score. The fact that there were repeated observations per individual did not affect the model much and therefore the cases are assumed to be independent.

The number of fixations on the banner was divided over two variables; the average number of fixations on a banner per participant and the deviation from this mean per participant, this to extract the effects of participants (i.e. some individuals may fixate more (or less) on banners than others to begin with). Page number was standardized and contrasted (one linear and one quadratic contrast have been employed).

One heavy leverage point was dropped from analysis, other outliers were kept. When all outliers would be dropped, the model would become even better (i.e. higher  $R^2$  and adjusted  $R^2$ ).

**Table 3.2** Multiple Regression model: Recognition scores

	Coef.	Std. Err.	Beta
Mean number of fixations on banner per individual	0.18	0.06	.14**
Deviation from mean number of fixations on banner per individual	0.15	0.02	.31***
Immediate questionnaire	0.62	0.18	.17***
Task-free condition	0.33	0.16	.09*
Page number Linear	-0.02	0.07	-.02 ( $p < 0.75$ )
Page number Squared	-0.16	0.07	-.11*
Age	0.07	0.03	.11*
Sex	-0.09	0.18	-.01 ( $p < 0.6$ )
Hours of internet use per day	0.04	0.04	0.04 ( $p < 0.45$ )
Constant	4.61	0.67	.***

N = 422;  $R^2 = 0.20$ ; Adjusted  $R^2 = 0.18$

Table 3.2 shows that the more fixations on the banner, the higher the likelihood that the banner will be recognized in a later stadium. When participants have to complete the recognition task immediately after the experiment there is a higher likelihood they will perform better than when they have to complete the task after a week. When a banner is shown during a free surfing task the likelihood that the banner will be recognized increases. The page number on which a banner has been placed follows an inverted U-shaped curve; there is a higher likelihood that banners at the middle of a sequence of pages are recognized compared to the first and last pages. The older participants are, the better their recognition scores are. There are no significant differences found between men's and women's recognition scores or for the number of hours participants spend online per day. Further no significant differences ( $F(9)=1.62$ ,  $p < 0.11$ ) in recognition scores between the different banners have been found.

### 3.1.2 Affective responses do not significantly differ between conditions

Affective responses to the advertised brands (e.g. brand attitude) are expected to be higher for banners which are unconsciously processed (because the mere exposure effect that increases the feeling of familiarity) and lower for banners which are consciously processed (because attention is expected to have a negative relationship with attitudes). The control condition is expected to be in between conscious and unconscious processing.

Whether a banner has been processed pre-attentive or attentive is difficult to determine because of individual differences and differences in the characteristics of the banner (Nelson & Loftus, 1980 in Wedel & Pieters, 2007). Earlier research does not agree upon the range of degrees for focal vision<sup>2</sup> or the minimum number of milliseconds needed for attentive processing. In this experiment it is assumed that banners are processed attentively when the maximum fixation length on the banner has a duration of at least 100 milliseconds (which is most common used), else it is assumed it is processed pre-attentively.

Conscious (i.e. attentive) processing of the banner was set to “1” if the banner was fixated for at least 100 milliseconds, the time the eye generally needs to process information, and to “0” if the banner was fixated for less than 100 milliseconds.

The construct brand attitude was measured by three questions (e.g. “I consider [brand] to be: ...” with a 9-point semantic differential scale anchored by “positive/negative”). Factor analysis confirmed that the covariances of the three questions naturally reproduced the construct. This factor ranged from -1 (bad) to 1 (good). The variable brand in consideration set was set to 1 when the brand was considered, and else to 0. Purchase intent scores ranged from 1 to 5, with 1 being very unlikely to purchase and 5 very likely to purchase. Brand favorability ranged from 1 to 3, with 1 being the best and 3 being the worst.

**Table 3.3** Descriptives: Affective measures

	<b>Conscious Banner Processing</b>	<b>Unconscious Banner Processing</b>	<b>Task (exposed condition)</b>	<b>Task-free (exposed condition)</b>	<b>Control Condition (not exposed to banners)</b>
	N=165	N=257	N=225	N=197	N=334
Brand attitude (ranging from -1; bad, to 1; good)	-0.01 (SE=0.07)	0.01 (SE=0.06)	0.01 (SE=0.05)	0.00 (SE=0.08)	-0.04 (SE=0.06)
Brand in consideration set (1=yes, 0=no)	0.09 (SE=0.02)	0.14 (SE=0.02)	0.12 (SE=0.02)	0.14 (SE=0.02)	0.12 (SE=0.02)
Purchase intent (ranging from 1; very unlikely, to 5; very likely)	2.08 (SE=0.07)	2.15 (SE=0.04)	2.13 (SE=0.04)	2.15 (SE=0.07)	2.12 (SE=0.05)
Brand favorability (ranging from 1; best, to 3; worst)	2.04 (SE=0.06)	2.04 (SE=0.03)	2.06 (SE=0.03)	1.98 (SE=0.06)	2.06 (SE=0.04)

<sup>2</sup> Banners can only be processed attentively when they are in the focal field of the eye, banners which are in the parafoveal field of the eye receive pre-attentive processing (Vanhuele et al, 2005).

Table 3.3 shows slight differences between the different conditions for each measure, however none of them is significantly different. Using different values for conscious processing (50, 150, 200 or 250 milliseconds) did not lead to significant differences either.

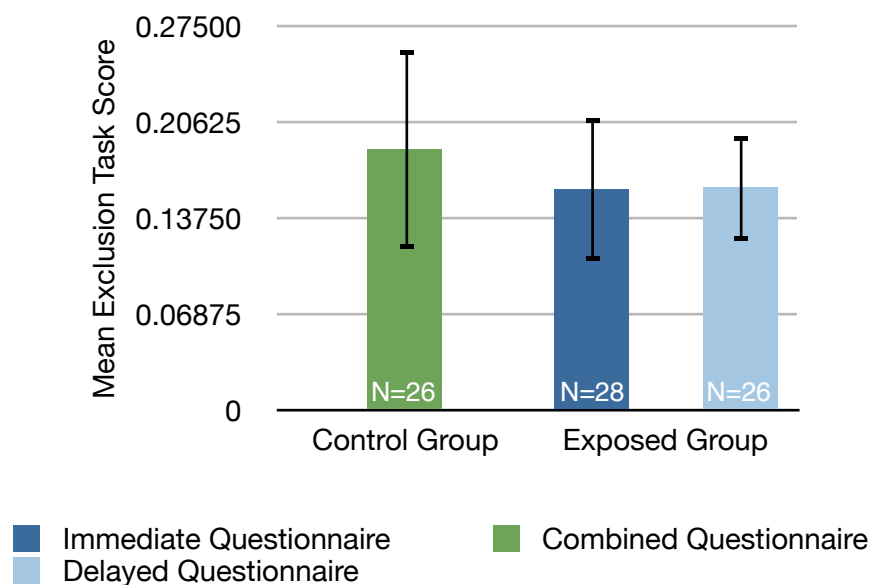
The scores indicate that;

- Brand attitude is neutral on average and is slightly better when a participant was exposed to a banner.
- The advertised brands are not often in the consideration set; brands are considered about 10% of the time .
- Purchase intention is on an average level; participants are not likely nor unlikely to purchase.
- Brand favorability scores show that the advertised brands are not considered to be better than their competitors or the other way around.

### 3.2 PDP scores do not significantly differ between conditions

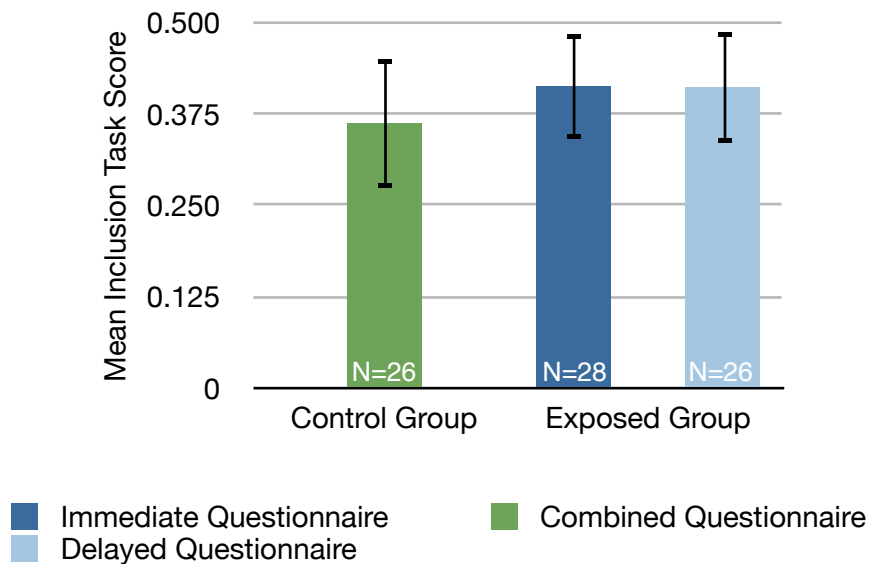
Process Dissociation Procedure scores are expected to be better for the exposed group; meaning lower scores for the exclusion task and higher scores for the inclusion task. No difference is expected between the immediate and the delayed questionnaire condition because implicit memory should last longer over time. The delayed- and immediate questionnaire conditions for the control group have been combined into one condition, because no implicit memories were created since there was no exposure to the banners and therefore time differences should not matter.

For the Process Dissociation Procedure, the scores range from 0 to 1, where for the exclusion task scores closer to 0 are better and for the inclusion task scores closer to 1 are better.



**Figure 3.5** Mean exclusion task scores for exposed vs control and immediate vs delayed questionnaire





**Figure 3.6** Mean inclusion task scores for exposed vs control and immediate vs delayed questionnaire

Figure 3.5 and 3.6 show that there are only small differences between the control and the exposed group, with the exposed group performing better for both the exclusion as the inclusion task, however these differences are not significant ( $p < 0.4$  for the exclusion task and  $p < 0.3$  for the inclusion task). The differences between participants who have filled out the questionnaire immediately after the experiment and participants who did so after a week are also non-significant ( $p < 1.0$ ), as expected since implicit memories are expected to last for at least one week.

### 3.3 Eye tracker data

The goal of analyzing the eye tracker data is to find differences in eye behavior between the exploratory search mode and the goal-directed search mode. It is expected that behavior in the free surfing condition is more focused and efficient and therefore it is expected that the eye and mouse movements are more clustered in smaller areas.

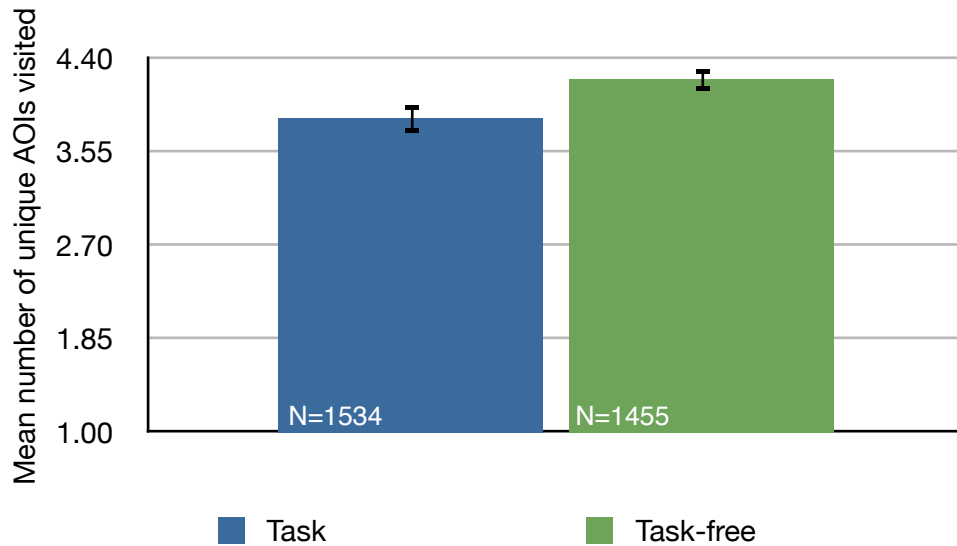
For the analysis of the eye tracker data itself only pages (i.e. urls) with a banner or a filler were selected. All these pages had the same lay-out. All other pages were dropped from analysis.<sup>3</sup>

#### 3.3.1 Number of Unique Areas of Interest (AOIs) higher for free surfing condition

The number of unique AOIs visited per page is expected to be higher for pages visited within the free surfing condition, this because surfing would be less efficient compared to the task condition.

All webpages have been divided into AOIs and per page a minimum of 1 and a maximum of 12 AOIs can be fixated upon (see figure 3.1).

<sup>3</sup> One heavy outlier was dropped from analysis since its leverage was much too high (ca. 19 times larger than  $3(k+1)/N$ , in which  $k$  = the number of predictors and  $N$  = the number of observations), all others were kept.



**Figure 3.7** Mean number of unique AOIs visited by task vs task-free condition

Figure 3.7 shows that the mean number of unique AOIs visited is higher for the free surfing condition, meaning that more unique areas of interest are visited within the free surfing condition compared to the task condition. This difference is significant ( $p < 0.001$ ).

To find out the weight of the condition on the number of unique AOIs visited, a multiple regression has been ran with time on page, the time stamp within the experiment, age, sex and the number of hours participants spend online per day as covariates. It is expected that the longer a person stays on a page, the more areas of interest are visited and the later in the experiment the less areas are visited because then it is known what information each area contains. All predictor variables have been standardized.

**Table 3.4** Multiple Regression: Number of Unique AOIs

	Coef.	Std. Err.	Beta
Task-free condition	0.27	0.04	.11***
Time on page	1.01	0.04	.41***
Time stamp	-0.45	0.04	-.18***
Age	-0.07	0.04	-0.03 ( $p < 0.1$ )
Sex	0.04	0.04	0.02 ( $p < 0.3$ )
Hours of internet use per day	0.03	0.04	0.01 ( $p < 0.5$ )
Constant	4.04	0.04	.***

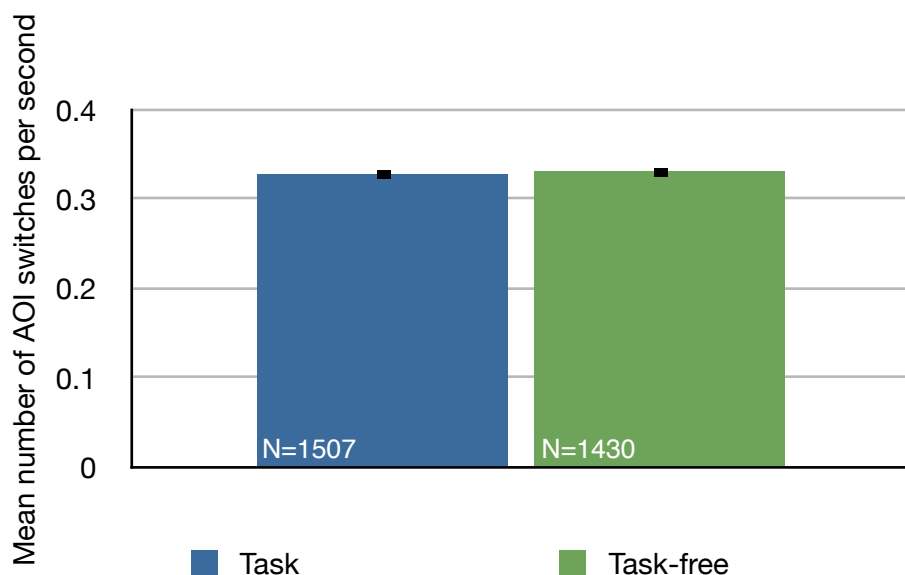
N = 2989;  $R^2 = 0.20$ ; Adjusted  $R^2 = 0.20$

Table 3.4 presents the results of the regression model. Significantly more AOIs were visited in the free surfing condition. There is an effect for time that participants spend on each page; the longer they stay on a webpage, the more AOIs they fixate upon (time on page accounts for most of the variance). Next there is an effect for time stamp; the later in the experiment the less AOIs are visited. Age, sex and the number of hours participants spend online do not significantly affect the number of unique AOIs visited.

### 3.3.2 Switching between AOIs does not significantly differ between conditions

The number of switches between AOIs (per page) is expected to be higher for pages visited within the free surfing condition, this because surfing would be less efficient compared to the task condition.

The number of switches between AOIs is determined by the sum of all switches from one AOI to another AOI within one page divided by the page visit time in seconds.



**Figure 3.8** Mean number of switches between AOIs per second by task vs task-free condition

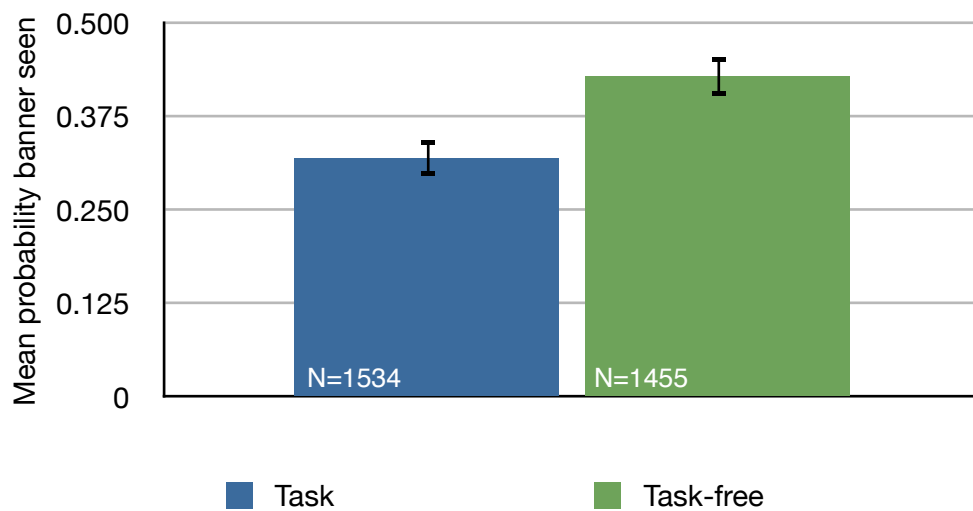
The mean number of switches between AOIs per second (Figure 3.8) are about as high for the free surfing condition and the task condition. The difference is indeed not significant ( $p < 0.4$ ).

Further differences in spatial density and the total number of fixations are regressed. The free surfing condition turns out to be a significant predictor (Beta < 0.06,  $p < 0.001$ ) in both cases, however the  $R^2$  and the Adjusted  $R^2$  are very low; 0.00 and 0.01 respectively.

### 3.3.3 Banners are more often viewed in the free surfing condition

The probability that a banner is fixated upon is expected to be higher for banners which are shown during the free surfing condition versus the task condition. This because attention is not so singularly focused in the free surfing condition and therefore banners can compete for attention.

If there is at least one fixation on the banner (or filler) then a score of “1” is denoted and else, if there are no fixations on the banner, a score of “0” is denoted.



**Figure 3.9** Mean probability banner seen by task vs task-free condition

Figure 3.9 shows that the mean number of times a banner is fixated upon is higher for the free surfing condition, meaning that banners are more often fixated upon in the free surfing condition compared to the task condition. This difference is significant ( $p < 0.001$ ). When controlling for time effects; since participants spend more time per page during the free surfing condition compared to the task condition, this effect remains.

To find out the weight of the condition on the probability that a banner receives at least one fixation, a logistic regression has been ran with time on page, the time stamp within the experiment, age, sex and the number of hours participants spend online per day as covariates. It is expected that the longer a person stays on a page, the higher the likelihood that a banner will receive (a) fixation(s) and the later in the experiment the less likely it is that a banner will receive (a) fixation(s) because then it is known what information each area contains. All predictor variables have been standardized.

**Table 3.6** Logistic regression: Banner fixated

	Coef.	Std. Err.	p
Task-free condition	0.27	0.04	***
Time on page	0.49	0.05	***
Time stamp	-0.14	0.04	**
Age	-0.09	0.04	*
Sex	-0.12	0.04	**
Hours of internet use per day	0.02	0.04	<0.6
Constant	-0.54	0.04	***

N = 2899; Pseudo R<sup>2</sup> = 0.05

Table 3.6 shows that the free surfing condition has a significant positive effect on the probability that a banner receives (a) fixation(s). Thus banners which are shown during the free surfing condition have a higher probability of being seen. Also the time participants spend on a page has a significant positive effect, indicating that the longer a participant stays on the page, the higher the likelihood that the banner will be seen. Time stamp, age and sex have significant negative effects, indicating that the further in the experiment, the lower the probability that banners will be seen; the older the participants are the lower the probability the banner will be seen and women have a higher probability that they will see banners compared to men. The number of hours participants spend online per day does not significantly affect the probability of banners being seen.

### **3.4 mouse movements do not correlate highly with eye movements and cannot predict search mode by itself**

Mouse and eye movements are expected to correlate highly, such that a mouse tracker could be used instead of an eye tracker in order to determine what areas are fixated. Mouse movements may -like eye movements- be more efficiently divided over the webpage when participants are in the task condition. Meaning that there would be less AOIs visited and less distance travelled.

During the experiment mouse and eye movements were measured with different programs and as a result their format is different; fixations were measured event-based, while mouse positions were generally captured during each computer cycle. The merging of these two data files into one file happened as follows; First the mouse time stamps were adjusted to match the same points in time as the eye time stamps; one event that appeared in both data files was chosen and the difference in time stamps for this event was subtracted from, or added to, the mouse time stamps depending on which recording started first. Second both mouse and eye data were merged into one file (with a marking of the kind of data; mouse or eye) and sorted by time stamp. Third the best possible match between eye and mouse time stamps was determined. The best match could be either the first mouse data case above or below an eye data case (because these are the most nearby in time stamp to begin with). The absolute difference between the first mouse time stamp above and the first mouse time stamp below the eye data case was calculated, and the case with the smallest difference was chosen to be the best match. After the data was matched, cases with differences greater than 100 milliseconds were dropped from analysis as these were considered to be too large differences.

#### **3.4.1 Mouse and eye coordinates are not highly correlated**

The X and Y coordinates of the eye- and mouse movements are not highly correlated; the correlation of the X coordinates of the mapped eye fixations and the mouse positions is 0.28 and the correlation of the Y coordinates of the mapped eye fixations and the mouse positions is 0.35. This implies that

mouse movements cannot reliably predict eye fixations; the probability that an individual is paying attention to another area is greater than the probability that she is fixating on the same area as the mouse is in.

When looking at the position of the mouse when the eye is fixating on the banner or filler, we can see that the difference in coordinates is on average 203 pixel for X ( $SD=252$  pixel) and -127 pixel for Y ( $SD= 174$  pixel). Meaning that the mouse is on average to the left of the banner or around the left half of the banner and on average a little higher positioned than the banner or around the upper half of the banner (i.e on the main content or the menu). When looking at the dispersion of the mouse coordinates (see figure 3.10) the main content and menu area are also the areas the mouse is most often situated.



Figure 3.10 Mouse positions heat map (over all pages)

### 3.4.2 Mouse movements cannot predict the mode the user is in

Even if mouse and eye movements are not correlated highly, the mouse might still be able to predict the mode by itself. There may be differences in clicks, the number of mouse movements, the distance

the mouse travels and the distribution of the mouse positions over the page. It is expected that the number of AOIs visited and the distance travelled are higher in case of the free surfing condition compared to the task condition.

The variables are constructed as follows; when either the X or the Y coordinate of the mouse changed this is counted as one additional change. The number of mouse clicks is self-explanatory. The distribution of mouse positions over the page could not be determined by looking at which AOIs were visited, because the AOIs did not start and end at the same coordinates for each page and the coordinates could not be exported from the eye tracker. Therefore, as an alternative, blocks of 100 x 100 pixels were made. The number of (unique) blocks visited ranged from 1 (as the mouse is always on the screen) to 192 (as the screen has a resolution of 1600 x 1200;  $16 * 12 = 192$ ). Finally, the distance the mouse travelled was calculated by the sum of the differences (in pixels) between the current and the previous mouse position per page.

It must be noted that only a part of the data (7357 of 93823 cases) has been used for analysis because of the imperfect synchronizing of the eye and mouse data and therefore the means are most likely a lot lower than they would be for the complete data. Nevertheless it should still be possible to spot differences between modes, because the number of remaining observations is still very large.

**Table 3.7** Descriptives: Predictors of mode

	<b>Task Condition</b> N=3851	<b>Task-free Condition</b> N=3506
Number of changes in X and Y coordinates	5.36 (SE=0.16)	5.47 (SE=0.15)
Number of mouse clicks	0.11 (SE=0.01)	0.10 (SE=0.01)
Number of 100 * 100 pixel blocks visited	2.48 (SE=0.07)	2.54 (SE=0.05)
Distance the mouse has traveled	373.30 (SE=10.03)	373.72 (SE=9.99)

When comparing the mean values of the above mentioned variables for the task and free surfing condition (table 3.7), we see that the differences between them are very small. T-tests confirmed that the differences are all non-significant. Therefore it is very likely that none of the variables in the table above can significantly predict the mode the user is in. A logistic regression confirmed this.

### 3.5 Summary

To summarize, the results show that there are only differences found for recognition and that all affective measures do not significantly differ between conditions. Recognition is best in the free surfing, exposed, immediate questionnaire condition and when the banner is placed in the middle of a sequence of pages within a task (e.g. page 2,3 or 4). Eye movements can predict the mode a user is in (e.g. whether a person is searching for specific information or is freely surfing) by counting the number of visited AOIs and by the time spent on the page. Mouse movements, on the other hand, cannot predict the mode the user is in and neither can they predict to what area the user is paying attention.

### 4. Discussion

This experiment confirms that there are responses to online advertisements beyond CTR. While there has only been one click on a banner in the whole experiment (i.e. a CTR of 0,002%), 42% of the banners have received at least one fixation<sup>4</sup>, which led to significantly higher recognition scores than the baseline guessing score of the control group. The highest recognition rates and the most fixations on a banner are obtained when a banner is shown during the free surfing condition. Recognition is best when it is measured shortly after the banner exposure, because the explicit memories of the banner decay pretty fast (after a week the recognition scores are not significantly higher than the baseline level). Therefore repetition is considered to be useful for advertisers. Also publishers can benefit from this by pointing out the advantages of repetition to advertisers.

Next to the cognitive measure; recognition, a set of affective measures (e.g. brand attitude and brand favorability) have been used. However no significant differences between the different kinds of processing (i.e. conscious, unconscious or no banner processing) or between the two different search modes have been found. This opposed to prior research which did find such differences. Also the process dissociation procedure shows that there are no significant differences between conscious and automatic influences between the control and the exposed group, despite slight differences in means in the direction of the expected effects.

The lack of an effect might be due to the more realistic setting of this study; the participants had more freedom and the webpages were more realistic (no mock-up website and/or content) and contained more distractions from other images and advertisements. Since the same measurements and the same kind of set-up as in previous research have been used, this cannot be the reason why no effect is found. When the banner advertisements themselves would have caused this, this implies that multiple real life banners do not receive affective responses. This would apply to unfamiliar brands at least. Familiar banners may receive affective responses after all, because they differ from unfamiliar brands in that they do not have to deal with the wear-in effect; new brands have to advertise more than

---

<sup>4</sup> The CTR may be biased because of the experimental setup, but when comparing with the average CTR of 0,09% (EMEA Region, Doubleclick report, 2009), the argument remains.



established ones to reach optimal effectiveness (Vakratsas & Ambler, 1999). However previous studies did find affective responses for fictitious, and therefore unfamiliar, brands. Therefore the cause is more likely to be found in the realistic setting.

Future research could investigate whether the more realistic setting causes the lack of an effect by comparing webpages with a main text, a lot of stimuli and a banner to a page with a main text and just one banner. If the effect turns out to only occur in controlled settings with little distractors than this is an important issue for advertising to know. Also it might be interesting to look into the effects of repetition on affective responses, because the differences in means may increase with repetition and become significant after all.

Further it must be noted that this study used the purchase funnel as a measure for banner effectiveness, with measures the closer to purchase the better. However campaigns do not always aim for purchase, they can also have different goals. What is important for one product category does not have to be important for another. For example explicit memory is more important for high involvement goods (i.e. expensive products which you do not buy everyday, e.g. a car or a computer) whereas implicit memory is more important for low involvement goods (i.e. inexpensive products that you buy on a daily basis or impulse purchases, e.g. chewing gum) (Shapiro & Krishnan, 2001). Therefore what is really most effective depends on the campaigns goal. However, this experiment shows that there are only significant differences for recognition, so no matter what the campaign's goal would be, banners shown during the exploratory search mode obtain the best effects.

Next to the search mode, the page number on which the banner is placed is important for the amount of attention a banner receives. Banners were best recognized when placed on page 2 or 3 within a task, which is not consistent with previous research. It was expected that banners would be more effective at the beginning and the end of a meaningful path (Wang & Day, 2007), but instead this study found that banners are more effective in the middle of a sequence of pages; the exact opposite. This implies that participants are focused at first, than quickly loose focus and than regain focus. Why this is the case is unclear. Future research may provide more insight into this matter. Also future research might look into banners on pages further than page five, because for now it is unclear how this curve evolves.

When it was clear in which search mode the best banner effectiveness is accomplished, mouse and eye data have been used to predict the mode the user is in. There are differences found in behavior for exploratory and goal-directed search modes. Within the exploratory search mode more unique AOIs are visited and there is a greater likelihood that banners receive fixations. Unfortunately the mouse tracker data is not as highly correlated with the eye tracker data as expected and by itself it does not provide useful information to predict the mode the user is in. This does not mean that a mouse tracker cannot predict behavior. This study was the first to look into this matter and used a separate eye- and mouse tracker which led to a suboptimal combining of data. The conclusions drawn from this data are based on a subset of the data that was well synchronized. To avoid dropping data and working with

gaps in your data, it is recommended to use a two-in-one tracker so that eye and mouse movements are optimally synchronized. By using a two-in-one tracker, it would be possible to track mouse movements to corresponding AOIs. In this study the latter was not possible because the AOIs did not always have the same coordinates and therefore could not be traced back to the mouse coordinates.

For now, the mouse tracker cannot be used on a large scale as a predictor of behavior and thereby banner effectiveness. Future research may develop different methods to extract information from mouse movement data that can predict behavior. Besides it should take into account that people develop different searching/scanning strategies based on the function and genre of the website (Josephson & Holmes, 2002 in Cooke, 2006), so maybe a mouse device can be effective for different kinds of websites as we only looked into one kind of website for this study. There may be differences between age and the number of hours of internet use for different participant groups. We only used one group; students. Using more diverse groups may lead to different results.

Nevertheless the mouse tracker cannot infer the mode the user is in, publishers can still use the information provided by this study by looking at the origin of the website visit; when someone visits the website via a search engine while looking for (a) specific keyword(s) it is very likely she will be in a goal-directed search mode, whereas someone who visits the website on a daily basis and directly types the url into the browser, is likely to be in an exploratory search mode. Future research may look into the origin of the website visit and the click through rate (as there is no other measure that can be used on a large scale so far). Furthermore the time spend on a webpage may indicate the mode a user is in, as participants spent more time per page in the free surfing condition compared to the task condition.

#### **4.1 Implications**

The insights gained from this study are useful for the advertising industry to bring online advertising to the next level by establishing a mutual understanding of how investments in advertising contribute to advertiser's objectives. It will drive marketing efficiency and effectiveness based on knowledge and understanding of the online advertising efforts by providing insights about how web visitors actually perceive and value online advertising and how long people remember exposures over time. Although these insights will not change the status quo (i.e. payment systems) overnight due to practicalities, it is a valuable step in really understanding (and therefore being able to address) online advertising effectiveness. It supports decision making on when (and when not) and how to advertise. Moreover, although this research is still exploratory, it provides the intuition that advertisers should assess the publishers audience, web site structure and banner placements in order to determine optimal configuration. There is a clear interest for the publisher as well. Being able to understand their web visitor behavior allows them to offer a more relevant proposition to the advertisers. This allows them to ask a higher price per impression. Finally, although it is not very likely that this methodology will be widely used across the industry in individual situations since this will be too costly, it allows for establishing industry-wide benchmarks. These can be used by supply and demand side to determine the fit for advertising.

## References

- Albert, W. (2002). Do web users actually look at ads? A case study of banner ads and eye-tracking technology. In *Proceedings of the Eleventh Annual Conference of the Usability Professionals' Association*.
- Anand, P., Holbrook, M.B., & Stephens, D. (1988). The formation of affective judgements: the cognitive-affective model versus the independence hypothesis. *Journal of Consumer Research*, 15 (December), 386-391.
- Benway, J. P., & Lane, D.M. (1998). Banner blindness: Web searchers often miss 'obvious' links, Internetworking. *ITG Newsletter*, 1(3).
- Biel, A.L. (1993). Ad research in the US: Hurdle or help? A brief review of the state of the art. *Admap* 329 (May): 27-29.
- Bornstein, R.F. (1989). Exposure and affect: overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, 106(2), pp. 265–289.
- Briggs, R., & Hollis, N. (1997). Advertising on the Web: Is there response before click-through? *Journal of Advertising Research*, 37, 33–45.
- Burke, M., Gorman, N., Nilsen, E., & Hornof, A. (2004). Banner ads hinder visual search and are forgotten. In *Proceeding CHI 2004*, ACM Press.
- Burke, M., Hornof, A., Nilsen, E., & Gorman, N. (2005) High-cost banner blindness; ads increase perceived workload, hinder visual search, and are forgotten. *ACM Transactions on Computer–Human Interaction*, 12(4), p. 423.
- Calisir, F., & Karaali, D. (2008). The impacts of banner location, banner content and navigation style on banner recognition. *Computers in Human Behavior* (24:2), pp 535-543.
- Chatterjee, P. (2008). Are unclicked ads wasted? Enduring effects of banner and pop-up ad exposure on brand memory and attitudes. *Journal of Electronic Commerce Research*, Volume 9, No. 1:51-61.
- Chattopadhyay, A., & Nedungadi, P. (1992). Does attitude toward the ad endure? The moderating effects of attention and delay. *Journal of Consumer Research*, Volume 19, No. 1:26-33.
- Chen, M.C., Anderson, J.R., & Sohn, M.H. (2001). What can a mouse cursor tell us more? Correlation of eye/mouse movements on web browsing. In *Proceedings of the Conference on Human Factors in*

*Computing Systems CHI 2001*, extended abstracts on Human factors in computing systems, Seattle, Washington, USA, April 2001.

Chtourou, M.S., Chandon, J.L., & Zollinger, M. (2002) Effect of price information and promotion on click-through rate for internet banners. *Journal of Euromarketing*, 11(2), pp. 23–40.

Cooke, L. (2006). Is the mouse a “poor man’s eye tracker”? In *Proceedings of the 53rd Annual Conference of the Society for Technical Communication* (pp. 252-255). Fairfax, VA: STC.

Cowen, L., Ball, L. J., & Delin, J. (2002). An eye-movement analysis of web-page usability. In X. Faulkner, J. Finlay, & F. Détienne (Eds.), *People and Computers XVI—Memorable yet Invisible: Proceedings of HCI 2002* (pp. 317-335). London: Springer-Verlag Ltd.

Dahlén, Micael (2001). Banner Advertisements through a New Lens. *Journal of Advertising Research*, Volume 41, No. 4:23-30.

Danaher, P. and Mullarkey, G. (2003, September). Factors affecting online advertising recall: A study of students. *Journal of Advertising Research*, Volume 43, No. 3:252-267.

DeSchepper, B., & Treisman, A. M. (1996). Visual memory for novel shapes: Implicit coding without attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 27–47.

DoubleClick 2009 research report: The Brand Value of Rich Media and Video Ads. Retrieved 04/15/11 from <http://www.google.com/doubleclick/pdfs/DoubleClick-07-2010-DoubleClick-Benchmarks-Report-2009-Year-in-Review-EMEA.pdf>.

Drèze, X., & Hussherr, F. (2003). Internet advertising: Is anybody watching. *Journal of Interactive Marketing*, Volume 17, No. 4:8-23.

Drunen, van A., Broek, van den E.L., Spink, A.J. & Heffelaar, T. (2009). Exploring workload and attention measurements with uLog mouse data. *Behavior Research Methods*, 41 (3), 868-875.

Egner, S.,Itti, L.,& Scheier, C.(2000). Comparing attention models with different types of behavior data. *Investigative Ophthalmology and Visual Science*, 41, S39.

Ewing, M., Napoli, J., & Du Plesses, E. (1999). Factors affecting in-market recall of food product advertising. *Journal of Advertising Research*, 39(4), 29–38.

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

Goodrich, K. (2007). Why the bad (or good) attitude? An exploration of banner advertising attention and brand attitude. In J. Mohr & R. Fisher (Eds.), *AMA Educators' Conference proceedings*, Vol.18 (pp.278–279). Chicago: American Marketing Association.

Goodrich, K. (2011). Anarchy of effects? Exploring attention to online advertising and multiple outcomes. *Psychology and Marketing*, Volume 28, Issue 4, pages 417–440, April 2011.

Haley, R.I. and Case, P.B. (1979). Testing thirteen attitude scales for agreement and brand discrimination. *Journal of Marketing*, vol. 43 (Fall 1979), pp. 20–32.

Heath, R.G., Brandt, D., & Nairn, A. (2006) Brand relationships – strengthened by emotion, weakened by attention. *Journal of Advertising Research*, 46, 4, pp. 410–419.

Internet Advertising Bureau (IAB) Online Advertising Effectiveness Study 1997: Beyond the Banner.

Janiszewski, C. (1990a). The influence of nonattended material on the processing of advertising claims. *Journal of Marketing Research*, August, 263-278.

Janiszewski, C. (1990b). The influence of print advertisement organization on affect toward a brand name. *Journal of Consumer Research*, 17 (June), 53-65.

Janiszewski, C. (1993). Preattentive mere exposure Effects. *Journal of Consumer Research*, 20 (December), 376-392.

Janiszewski, C. (1998). The influence of display characteristics on visual exploratory search behavior. *Journal of Consumer Research*, 25 (3), 290-301.

Jaworska, J. & Sydow, M. (2008). Behavioural targeting in online advertising: an empirical study. *Web Information Systems Engineering - WISE 2008*.

John, B.E., & Marks, S.J. (1997). Tracking the effectiveness of usability evaluation methods. *Behavior and Information Technology* 16: 188-203.

Josephson, S., & Holmes, M. E. (2002). Visual attention to repeated Internet images: Testing the scanpath theory on the World Wide Web. *Proceedings of the Eye Tracking Research & Applications Symposium*, New York, NY: ACM, 43-49.

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

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

Klever, A. (2009). *Behavioral targeting: an online analysis for efficient media planning?* Hamburg: Diplomica Verlag GmbH.

Kunst-Wilson, W. R., & Zajonc, R. B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science*, 207, 557-558.

Lapa, C. (2007). Using eye tracking to understand banner blindness and improve website design. Unpublished master 's thesis, Golisano College of Computing and Information Sciences, Rochester Institute of Technology. <http://hdl.handle.net/1850/4768>

Lewicki, P., Hill, T. & Czyzewska, M. (1992). Non conscious acquisition of information. *American Psychologist* 47:796–801. [VGH]

MacKenzie, S.B., Lutz, R.J., & Belch, G.E. (1986). The role of attitude toward the ad as a mediator of advertising effectiveness: A test of competing explanations. *Journal of Marketing Research*, 23 (May), 130–143.

Mandler, G., Nakamura, Y., & Shebo Van Zandt, B. J. (1987). Nonspecific effects of exposure on stimuli that cannot be recognized. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 646-648.

Nelson, W. W., & Loftus, G. R. (1980). The functional visual field during picture viewing. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 391-399.

Nielsen, J., & Tahir, M. (2002). *Homepage usability: 50 Websites deconstructed*. New Riders Publishing, Indianapolis, IN.

Novak, T.P., & Hoffman, D.L. (2000). Advertising and pricing models for the web. In *Internet Publishing and Beyond: The Economics of Digital Information and Intellectual Property*, Deborah Hurley, Brian Kahin and Hal Varian, eds. Cambridge: MIT Press.

Pagendarm, M., & Schaumburg, H. (2001). Why are users banner-blind? The impact of navigation style on the perception of web banners. *Journal of Digital Information*, volume 2, No 1. Retrieved 4/21/11: <http://journals.tdl.org/jodi/article/viewArticle/36/38>.

Poole, A., Ball, L. J., & Phillips, P. (2004). In search of salience: A response time and eye movement analysis of bookmark recognition. In S. Fincher, P. Markopolous, D. Moore, & R. Ruddle (Eds.), *People and Computers XVIII-Design for Life: Proceedings of HCI 2004*. London: Springer-Verlag Ltd.

Poole, A., & Ball, L. J. (2005). Eye tracking in human-computer interaction and usability research: current status and future prospects. In Ghaoui, C. (Ed.), *Encyclopedia of Human Computer Interaction*, Idea Group.

Rettie, R., Grandcolas, U. & McNeil, C. (2004). *Post Impressions: Internet Advertising without Click-through*. Kingston Business School, Kingston University.

Richardson-Klavehn, A., & Bjork, R.A. (1988). Measures of Memory. *Annual Review of Psychology*, 39, 474-543.

Rodden K, Fu X, Aula A, Spiro I. (2008). Eye-mouse coordination patterns on web search results pages. *Proceedings of the CHI '08 Extended Abstracts on Human Factors in Computing Systems*. 2008. ACM Press, New York, pp. 2997–3002.

Rosenkrans, G. (2009). The creativeness and effectiveness of online interactive rich media advertising. *Journal of Interactive Advertising*, 9(2).

Russo, J.E. (2010). Eye fixations as a process trace. In M. Schulte-Mecklenbeck, A. Kuehberger and R. Ranyard (Eds.), *A handbook of process tracing methods for decision research*. London: Psychology Press.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 501–508.

Shapiro, S. (1999). When an ad's Influence is beyond our conscious control: Perceptual and conceptual fluency effects caused by incidental ad exposure. *Journal of Consumer Research*, 26, 16–36.

Shapiro, S., & MacInnis, D. J. (1992). Mapping the relationship between preattentive processing and attitude. In J. F. J. Sherry & B. Sternthal (Eds.), *Advances in Consumer Research* (Vol. 19, 505–513). Chicago: Association for Consumer Research.

Shapiro,S.,& Krishnan,H.S.(2001). Memory-based measures for assessing advertising effects: A comparison of explicit and implicit memory effects. *Journal of Advertising*, 30,1–14.

Singh, S., Rothschild, M., & Churchill, G. (1988). Recognition versus recall as measures of television commercial forgetting. *Journal of Marketing Research* 25, 1: 72-80.

Tellis, G. J. (1998). *Advertising and sales promotion strategy*. Addison-Wesley Educational Publishers, Inc.

Theuner, G., Pischke, K., & Bley, T. (2008). Analysis of Advertising Effectiveness with Eye Tracking. In A.J. Spink, M.R. Ballintijn, N.D. Bogers, F. Grieco, L.W.S. Loijens, L.P.J.J. Noldus, et al. (Eds.), *Proceedings of Measuring Behavior 2008*. Maastricht, The Netherlands, August 26-29, 2008.

Tulving, E., Schacter D. L., & Stark. H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 336-342.

Vakratsas, D. & Ambler, T. (1999). How advertising works: what do we really know? *Journal of Marketing*, 63 (January), 26–43.

Vanhuele M., Courbet D., Denis S., Lavigne F., & Borde A. (2005). The effectiveness of non-focal exposure to web banners ads. *Finanza Marketing e Produzione*, XXIII, 3, pp. 7-12.

Wang, J.C., and Day, R.F. (2007). The Effects of Attention Inertia on Advertisements on the WWW. *Computers in Human Behavior* (23:2007), pp 1390-1407.

Wedel, M., & Pieters, R. (2007). A review of eye-tracking research in marketing. In N. Malhotra (Ed.), *Review of marketing research, Volume 4* (pp. 123–146). New York: M. E. Sharpe Inc.

Yoo,C.Y. (2005). Preattentive processing of Web advertising. Dissertation, University of Texas at Austin.


Yoo, C. Y. (2007). *Preattentive processing of Web advertising*. New York: Cambria Press.

Yoo, C. Y. (2008). Unconscious processing of Web advertising: Effects on implicit memory, attitude toward the brand, and consideration set. *Journal of Interactive Marketing*, 22(2), 2-18.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151–175.



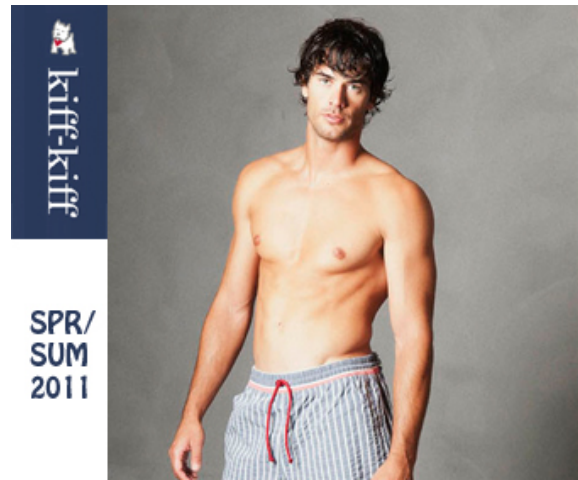
Appendix 1 - Experimental banners

Experimental banner:	Alternative for forced choice in questionnaire:
 <p><b>KISS MY GLASS</b></p> <p>MOD. SCIENTIST COL. RED</p> <p><b>GlaSSiNG</b></p>	 <p><b>KISS MY GLASS</b></p> <p>MOD. SCIENTIST COL. EMERY BLUE</p> <p><b>GlaSSiNG</b></p>
 <p><b>uniek in brillen</b></p> <p><b>ALAIN AFFLELOU</b></p>	 <p><b>voor hem en voor haar</b></p> <p><b>ALAIN AFFLELOU</b></p>
<p>UNLIMITED OFFERS. YOU WON'T PUT YOUR MOBILE DOWN. </p> 	<p>UNLIMITED OFFERS. YOU WON'T PUT YOUR MOBILE DOWN. </p> 

Experimental banner:



Alternative for forced choice in questionnaire:



Experimental banner:



100% recyclebaar

Alternative for forced choice in questionnaire:



met ECO Quality!



The New U-line Available Now!



ACOLFO DOMINGUEZ



New spring/summer collection



ACOLFO DOMINGUEZ

Wereldwijd bellen

میدیتل  
Méditel



Whenever, Wherever

میدیتل  
Méditel

**Experimental banner:**



**Alternative for forced choice in questionnaire:**



Appendix 2 - Filler ads

	  <p>Nu ook op je mobiel!</p>
<p><b>Uw advertentie hier?</b></p>  <p><b>Uw advertentie levensgroot op deze internet pagina!</b>          Zeer goed door Google gevonden!          Klik hier voor meer informatie &gt;&gt;</p>	 <p>3x <b>klikken</b></p> <p><b>Kijken</b></p> <p><b>Kiezen</b></p> <p><b>Kopen!</b></p>

## Appendix 3 - Questionnaire

Distractor Tasks
What <b>did you like</b> about the <b>Kliknieuws website</b> ? <i>Think of information, navigation, design, usability etc.</i>
What <b>didn't you like</b> about the <b>Kliknieuws website</b> ? <i>Think of information, navigation, design, usability etc</i>
PDP exclusion task
In a moment you will get a couple of word stems which you have to complete. The idea is to complete the word stem with the <b>first word that comes to mind of at least 4 letters</b> (in total). Besides the word may <b>not fit the advertisements</b> (banners) which were shown during the experiment.  Please fill out the <b>complete words</b> .  <b>Example:</b> You have seen an advertisement for a computer and you get the word stem <b>co</b> __ Then you may <b>not</b> complete the word stem with: <b>computer</b> . But you may complete the word stem with for example: <b>coordinate, container or cocoon</b>
Complete the following word stem with the <b>first word that comes to mind of at least 4 letters</b> that <b>does not fit the advertisements</b> (banners) which were shown during the experiment.  <b>gla</b> __
--- 5 more word stem completion tasks ---
PDP inclusion task
Complete the following word stems with <b>words that do fit the advertisements</b> (banners) which were shown during the experiment.  <b>Example:</b> You have seen an advertisement for a computer and you get the word stem <b>co</b> __ Then the idea is that you complete the word stem with: <b>computer</b> .
Complete the following word stem with the <b>first word that comes to mind of at least 4 letters</b> that <b>does fit the advertisements</b> (banners) which were shown during the experiment.  <b>te</b> __
--- 5 more word stem completion tasks ---
Purchase Funnel

### Brand awareness

Are you familiar with the **sunglasses brand Glassing**?

- Yes, I have heard of it / have seen it
- No, I don't know it / have not seen it
- I'm not sure

### Brand attitude

I consider **sunglasses brand Glassing** to be...

- Scale from negative to positive
- Scale from bad to good
- Scale from unfavorable to favorable

### Consideration set

Imagine you are looking to buy a pair of **sunglasses**. Which of following brands would you consider?

- Giarre
- Rudi Kellner
- Shade Station
- United Shades
- Glassing
- None of above

--- answer options are in a random order ---

### Purchase intent

Image you are looking to buy a pair of **sunglasses**. How likely is it that you would buy a pair of **sunglasses from Glassing**?

- Scale from very unlikely to very likely

### Brand favorability

Make a top 3 of following brands:

- United Shades
- Glassing
- Rudi Kellner

--- answer options are in a random order ---

### Ad awareness

Which of the following two banners have you seen during the experiment?  
When you do not know which one was shown, select a banner anyway.



How **certain** are you that you have selected the **correct banner**?

- 50%
- 60%
- 70%
- 80%
- 90%
- 100%

--- The purchase funnel questions are repeated for each banner ---

### Control questions

Were there brands that you already knew (if only by name) before this experiment? If yes, specify:

Were you familiar with Kliknieuws before this experiment?

- No, never heard of
- Yes, heard of but never visited the website
- Yes, I visit the website occasionally
- Yes, I visit the website regularly

How many **hours per day** on average do you use the **Internet**?

- .... hours

Sex:

- Male
- Female

Age:

- ... years