Visibility

Poster panel visibility for drivers and passengers: a first look

route
outdoor media analytics

Biographies

Dr Paul Barber was employed at Birkbeck in the University of London where for over thirty years he taught courses on perception, psychological statistics and experimental design, computing and ergonomics. He was head of Birkbeck's Department of Psychology from 1988 to 1992. He is now Emeritus Reader in Psychology at Birkbeck. Dr Barber was research supervisor to sixteen PhD students, his own doctoral research being on visual input processes. He is an Associate Fellow of the British Psychology Society and is a Chartered Psychologist. He was the author/co-author of a number of psychology textbooks and many research papers in refereed scientific journals; he was Psychology Editor for the journal Ergonomics for over ten years. He has been a consultant for Postar since 1995, closely identified with its programme of research on poster panel visibility.

Simon Cooper has a BSc in Cognitive Science from Birkbeck and has worked in outdoor research for over twenty years. He was the NOP project manager and architect for the Postar research in the UK in 1996 and before that worked on the original OSCAR research. This connection between academia and commercial research allowed for the integration of eye-tracking studies from Birkbeck to play a role in the development of visibility measurement in outdoor. As part of this early project he co-authored the Postar roadside visibility models with Dr Barber and co-designed all the visibility research between 1995 and 2007.

Contents

Foreword

Between 1995 and 2010, Postar commissioned a series of studies into the visibility of various outdoor advertising formats. These were undertaken by Birkbeck College at the University of London. Over time a considerable body of work was created that addressed the importance of size, distance, angle of presentation and so on. The studies also considered what it was like to be a car driver or passenger, a pedestrian or a passenger on board a train or bus. They looked at illumination and environment. In recent years, the work has progressed to begin to understand the effect of movement, such as that presented by digital screens.

A full list of the various studies can be found in the appendix.

Everything has to start somewhere and we thought it would be useful to those with an interest in the topic if we were to publish the research paper that accompanied the very first study. The fieldwork was conducted in 1995 and 1996. This seems like an age away and certainly much has changed since then. A great deal has since been built on these early foundations.

If you wish to understand the genesis of visibility research for out-of-home media, there can be no better place to start than at the beginning.

James Whitmore
Managing Director
Route

February 2013

Abstract

Estimating the visibility of poster panels for drivers and passengers

Paul Barber *and* **Simon Cooper**
Department of Psychology, Birkbeck College
Consultant to Postar

The visibility of poster panels for car drivers and passengers was assessed. This was done by recording eye movements while photographs of scenes were viewed for a fixed interval of six seconds. This interval was based partly on the duration of a typical drive past a poster panel site, taking account of the need to allow a sensible opportunity to register and inspect each scene as if driving, and to avoid boredom while sustaining interest in the driving task. The incidence of fixations on target panels was measured and used to provide a visibility score or *hit rate* for each panel depicted. Mean hit rates were obtained for three panel sizes (6, 48 and 96 sheets). The scenes used were from three road *environments* – Arterial, Residential and Shopping – selected as representative of where panels are placed in the UK outdoor advertising environment, and considered to vary in their degree of visual clutter.

A full factorial combination of panel size and environment was not possible because of how panels are and can be located (e.g., small formats are typically situated very close to the kerb-side while large formats tend to be at much greater offsets). The distribution of panels of the three sizes investigated was markedly unbalanced and this affects what can be concluded about comparisons between panel sizes and other factors. The results showed a general trend for hit rate to increase with panel size. Furthermore hit rate tended to be rather lower for Shopping scenes than Arterial or Residential scenes. Hit rate was also found to vary with panel *eccentricity*, decreasing as offset from the roadside increased although the incomplete nature of the experimental design meant that the trend was not comprehensively established.

The subjects participating in the research were asked to adopt the role of Driver or Passenger as they viewed the scenes. Hit rates for Passengers were higher than for Drivers, minimally for the small panels but to a greater extent for the larger two panel sizes. Additional research is required if an account incorporating the visibility of poster panels on the part of Pedestrians.

The manner in which hit rate accumulates as time since the scene is displayed was analysed. Fixations directed at poster panels tended to occur relatively early in the display interval.

A preliminary attempt was made to identify the possible functional regularity between hit rate and panel size. This was of limited success and further research is needed prior to the development of a visibility model.

Introduction

Preliminary considerations: Visibility and methodology

The efficacy of an advertising poster panel is quantified by the size of its audience, however, this is an imperfect measure unless properly qualified. It may be a relatively straightforward matter to estimate the frequency with which an advertising poster panel is in view of its audience; that is, how many people come within visual reach of the panel. On the other hand not all of those individuals see the panel, so the recorded frequency arguably needs to be adjusted to take this into account. This signals the need for an appropriate way of achieving the correction, and this defines the purpose of the present study. It is proposed here that the correction or adjustment should be done by determining and applying a measure of the panel's *visibility* (the rationale for the adoption of this concept is discussed in Appendix A).

The general aim of the study therefore is to assess the visibility of poster panels for drivers. This entails the collection of evidence regarding what we are terming the *visibility* of roadside poster panels, and from these foundations to develop a model for estimating the varying efficacy of poster panels. The process should ignore poster content if possible but should ostensibly take into account the eccentricity, luminance, angular size and setting of the panel. To achieve the aim of the study we need to derive an operational specification of visibility; to develop a principled and applicable account of the visibility of roadside poster panels; and allied to these objectives, we require a methodological basis for the acquisition of data. We begin with the question of research method.

Research objectives and choice of method

The study pursues a relatively novel line of enquiry though this is related to the broader question of which objects catch the attention of a driver, a topic that has been of interest to various research communities and their work supplies useful starting points for the present investigation. In particular these investigations demonstrate a range of methodological options for addressing the general issue as well as the specific version which is the focus of our research. A brief survey of these options follows, illustrating the methods available and the reasoning leading to the eventual choice of method – that of eye movement recording. This is a very attractive option because it directly reveals the trajectory of the eyes as the driver views the scene ahead. It has the potential added bonus of logging when a fixation on an object of interest (e.g., a poster panel) takes place, how many times that happens, and for how long.

The most full-blown version of this approach would be to record the eye behaviour of drivers in real-world settings as they drive a route or routes featuring poster panels. However, much of the time might well be spent driving without posters in view so a great deal of wasted time and eye movement recording would be the result. Moreover it would require a specially instrumented car equipped with eye-tracking equipment, and it would generally be a very expensive undertaking. The processing and analysis of the results would also be immensely demanding. Another possibility is to accept a decrease in the ecological fidelity of the method by using a laboratory-based approximation to what a driver encounters – including poster panels. For example, a series of images depicting driving scenes would need to be presented to an observer/driver while his or her eye movements were being recorded. Two versions of this arrangement are conceivable in principle at least; in the first the image could be provided by film (or video) clips obtained from a moving vehicle, in the second case by a set of still photographs. Although developments in eye movement recording and the requisite computing facilities will no doubt make a dynamic alternative easier to implement (and more affordable), this will only be a serious possibility in a few years' time – a proviso that applies with even greater force to the "full-blown" option – that is, the dynamic in-car recording method.

A different set of options is furnished by tasks that do not rely on the deployment of eye movement recording equipment. For instance, one approach would be to ask observers to find "targets" (defined as poster panels) in photographs of road scenes, using speed of response as an indicator of visibility. The drawback of this method is that the observer would be required consciously to search for the target(s). This is not what the majority of drivers do. They generally strive to use their eyes to aid the task of driving safely, not to search consciously for something unrelated to driving – e.g., poster panels. This is not to say that a directed search task could not supply useful adjunct data on poster panel visibility, but this would be a matter for empirical confirmation; for example, by showing that results from a directed search task were consistent with data from a task with accepted validity with respect to visibility measurement – such as eye movement recording. A second approach would be to ask observers to nominate the objects in a scene that were subjectively most visually prominent. A combination of qualitative and quantitative analyses would need to be applied to assess the prominence (i.e., visibility) of poster panels relative to one another. The method would be flexible and quick to yield data, but to be convincing from a commercial/practical perspective it would also need to produce results in agreement with those from a method with unimpeachable validity (undoubtedly again this would be eye movement recording). Interestingly there are a handful of reports in this vein: perhaps the most important and best known of these studies is by Mackworth and Morandi (1967) who compared the frequency of visual fixations on regions of pictures and verbal comparisons of the importance of those regions. Regions rated as of high informativeness were those that were fixated most frequently. Interestingly the fixation sequence for the eye movement group was established quickly (and without fully scanning the image). The implementation and use of the ratings task is quick and relatively simple to achieve relative to the eye movement task. It could provide a useful model for future studies, especially when indicative results are required in a short time-frame, but its applicability for the present study is limited in the absence of confirmatory data from an eye-tracking task.

It will be appreciated from the foregoing that the decision to use eye movement recording to ascertain what people look at when driving was not a matter for lengthy debate so long as appropriate equipment was available and affordable. Eye movement recording is the method that by definition enables a researcher to assess what an observer is looking at. Notwithstanding there are issues of validity (e.g., does having one's eye movements monitored affect one's visual behaviour?) and reliability (e.g., does the recording equipment render the data accurately and consistently?). Both problems have declined in significance as the recording devices have become less invasive and their operation more precise. For example headsets bearing the sensors (as used in this study) have become lighter and more comfortable and further developments may be anticipated as miniaturisation of computing and related components proceeds.

Although eye movement recording has a history reaching back to the end of the 19th century is was not until the 1960s and 1970s that interest in the technique and its potential showed significant growth, fuelled by the increasing provision of the technical platforms needed. This movement was enhanced in the 1980s by the advent of cheap and flexible computer systems that interface with devices to register eye behaviours. The bulk of the early research was on the psychology of reading but important steps were also achieved in the description and understanding of skills such as driving. The technical requirements to handle the data acquisition and analysis aspects of these two contrasting topics are inevitably quite different, with technological developments for some time favouring advances in research on skilled reading. In the event a task intermediate between the two was adopted for the purpose of the present research. This employed the two dimensional format of a computer screen (as used in reading research) for the presentation of static roadside scenes as viewed by a driver.

The implementation of the research technique for the purpose of the study required answers to a range of questions which will be dealt with in further detail in the Method section of this report.

The eye-mind hypothesis and our choice of research method

The key rationale for recording eye movements is that what a person's eyes are fixated on is indicative of what he or she is acquiring information about. A "strong" version of this eye-mind hypothesis was formulated by Just and Carpenter (1980) in relation to their theory of reading. They expressed this in the form: "there is no appreciable lag between what is fixated and what is processed". So in their research topic of reading if a subject looks at a letter or word he or she can be considered to "think about" – or cognitively process that letter or word. This information process continues until the fixation is ended and a new letter or word is fixated.

The eye-mind hypothesis generalizes to the viewing of other entities, including pictures and scenes. We do not need to adopt such a restrictive point of view and indeed this would be wise in the light of evidence of "covert attention". This refers to attention being paid to objects which are not being fixated and which may for instance steer future fixations via information acquired from peripheral vision. It is nevertheless sensible to assume that information is most likely to be extracted from locations that are fixated. Indeed the evidence from eye movement studies is that people do look at regions judged to be of potential interest in a picture. They also tend to revisit those locations visually and to do this in a fairly ordered fashion, reflecting a sequence of noteworthy features or "scanpath" (Noton and Stark, 1971). In the previously cited study by Mackworth and Morandi (1967), two separate groups of observers were tested, one asked to look at pictures and report what they judged to be most informative, the other asked just to look at the pictures while their eye movements were noted. What the two groups judged as informative or simply looked at was in good agreement. This is an important conclusion from a methodological point of view but also as evidence supporting the general eye-mind hypothesis.

Measuring visibility

As noted above eye movement recording was the methodology chosen to assess each viewer's visual behaviour in scenes containing poster panels. From the raw data files a record could be compiled of how many times each observer *looked at* or *fixated* a particular panel. If the eyes come to rest on the panel a fixation is considered to have occurred and a "hit" is registered. The results for a sample of observers may then be aggregated, thereby providing a score for a particular panel, defined by the proportion of observers scoring at least one hit on that panel. This is referred to as the "visibility hit rate" (or just "hit rate") for the panel. Averaging over a set of panels with a common property (say, size) provides the visibility hit rate for that set.

It should be noted that multiple hits by one observer contribute only once to the hit rate as defined. Consider a single panel viewed by 50 observers, of whom 20 fixate the panel at least once: the hit rate for the panel is 20/50 = 40%. Even if they all fixate it twice, by the definition the hit rate is still 20/50 = 40%. Similarly if only one person fixates the panel, but does so 20 times, the hit rate is 1/50 = 2%. An extreme position such as this last case is quite atypical, but it may be informative to compile a score based on the total number of hits, thus allowing multiple fixations by one or more observers to count towards that score.

A score incorporating multiple hits would clearly constitute a flawed measure of visibility, but it has potential value as a measure of the persisting visual effect of a panel beyond the first fixation that is directed at it. Other measures such as the duration of the first fixation on the panel, the latency of the first fixation (i.e., the time from display onset) and the duration of all fixations on the panel are also informative about visibility, but about more besides. Their shared shortcoming is that they probably reflect the panel's content (execution) more than does a simple hit rate measure.

Literature search and review

In order to evaluate the possibility of developing an empirically based visibility model for roadside poster billboards a search was made of recent scientific literature for similar experimental work. It became apparent that there was no research on poster billboards, but there were various studies on driver behaviour and road signs. In any study that could be undertaken and in reviewing previous work in this field, some assumptions must be made: first, what a driver looks at is what he/she attends to. Second, drivers tend not to search their visual environment, they just notice features in it.

Early studies on perception (e.g., Sperling 1960) researched what could be perceived in a brief (sub-second) exposure of visual information. It was found that up to nine items (e.g., characters or digits) could be perceived and reported by the subjects under test. The displays used were generally presented to the centre of the visual field. Further work on perception showed that performance depended on the location on the retina on which the task information was projected. For example, Sanders (1963) showed that performance in acquiring visual information was related to the position of that information in the visual field. It is evident that if an object is in the periphery then the eye must move (or make a "saccade") towards the object before a useful fixation can take place.

Some early work on car driver perception was carried out by Johansson and Rumar (1966) to quantify the ability of car drivers to get information from road signs – by finding the percentage of drivers who noted a given road sign. Their experiment entailed the placing of five different traffic signs of varying urgency at the road side before a blind corner, behind which the experimenters lay in wait – with the aid of the police – to stop and question passing drivers. Their results showed that even with observers instructed to spot road signs, their performance was only 90% correct. The results obtained by questioning drivers showed that an important sign (pre-warning for a speed limit zone) was registered by only 78%; and at the other extreme – a nominally less important sign (pedestrian crossing pre-warning) was registered by a mere 17%. The average recorded by drivers for all five signs used was 47%. The authors concluded (p62) that: "..., *it is not an exception but rather a rule that drivers overlook traffic signs.*"

A substantial body of research has accumulated on object conspicuity, a topic closely allied to the concerns of the present study. The concept was proposed and first investigated by Engel (1971) who suggested the conspicuity area of an object should be defined as the area surrounding it within which it could be seen (with a critical probability specified for the investigation). Engel's research was laboratory-based and the step into a field setting was taken by an Australian group of researchers (Cole, Hughes and Jenkins) whose chief interest was in traffic engineering and safety. Some of their work – an important source of evidence and ideas – is described in the next few paragraphs.

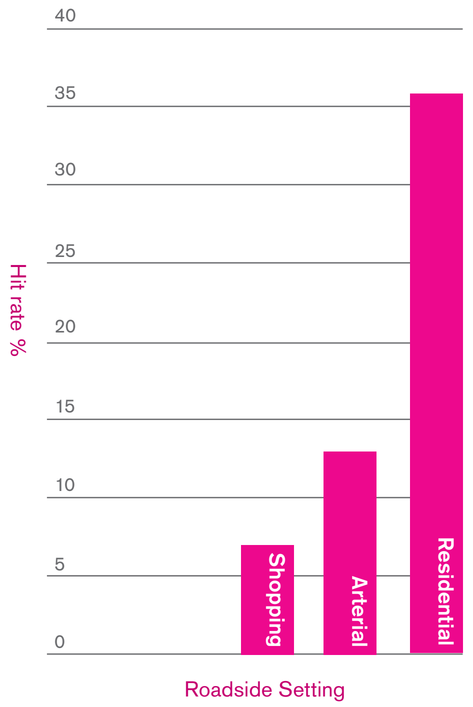
A study undertaken by Cole and Jenkins (1982) looked at the effect of complex backgrounds on target conspicuity. The results indicated that the size of a detectable object was related to the variability of the size of background objects; if the background objects have great variability of size the target object needs to have a substantial difference in size to be noticeable. Another result of this study was in demonstrating the importance of luminance: it was found that the variability of background luminance had little or no effect on target detection, which instead depended primarily on the average background luminance.

In a follow-up study Cole and Hughes (1984) undertook a more detailed examination of object conspicuity in a real world environment. They conducted a field trial in a Melbourne suburb in the course of which a group of subjects were required to drive along a defined route on which disc targets had been placed and asked to report what attracted their attention. The discs had been placed where conventional traffic signs could be expected. The subject "runs" were videotaped for later analysis of subjects' verbal reports and the distances at which reports of discs were made.

The fundamental measure of their study was the reported frequency of disc targets, which they termed the "hit rate". The results showed that object conspicuity was not strongly dependent on reflectance or size (in the limited range of disc diameters of 70cm, 50cm and 30cm). Cole and Hughes asserted (p306) that: "...*because the study was a field trial and the observers moved continuously through their visual environment, this simple analysis of the effect of target size may be misleading. It may be more appropriate to consider the projected angular size of the target discs at the time they were reported.*"

In this study Cole and Hughes also looked at the effect of three varieties of road environment: the character of these three road types (arterial, shopping and residential) was shown to have a highly significant effect on hit rate (see Figure 1) in a transposed histogram adapted from the Cole and Hughes report.

Figure 1: Hit rate (% registered) of disc targets as a function of road type (after Cole and Hughes, 1984)



The most important determinant of hit rate they reported in this study was the angle at which the object was displaced away from the line of sight. They concluded (p310) that: *"Our result suggests that in order to achieve conspicuity, the designer is better advised to locate the target where it will have a small eccentricity to the observer's line of sight rather than increase the size of the target."*

Another relevant study was performed by Unema and Rotting (1992) investigating the duration of the fixations made by drivers under varying mental workloads. In an experiment on open roads, using 20 bus drivers and 12 car drivers in Maastricht in the Netherlands, they attempted to measure fixation duration under varying traffic situations on set routes. Using an eye mark recorder and parallel video footage, simultaneous identification of fixations and areas of interest was possible. Because the manual frame by frame analysis was so laborious they only analysed selected sections of footage for areas of interest. However, the eye mark recorder was used to examine the number and duration of fixations, (though not points of interest) for the whole of the driven routes. The results showed that the mean fixation duration was between 150 and 450ms. Most drivers made two or three fixations per second depending on the complexity of the situation he/she was in.

In their review of the conspicuity of road signs for drivers, Cole and Hughes (1992) suggested that some of the important determinants of conspicuity are eccentricity, background complexity, and contrast. They also suggested several other factors, including colour and boldness of the internal structure of the object. For present purposes, attributes of the poster itself – such as content (including its colour properties) – are the responsibility of the advertiser not the site designers or owners and are not examined in the present study. Cole and Hughes stated that the conspicuity of an object *"might be defined as the property that leads to a target object having a high probability of being seen within a very short time. An object that does not have a high probability of being seen or noticed, or is only seen after a lengthy period of search, cannot really be described as conspicuous."*

Although in their review of previous studies Cole and Hughes concluded that size does not immediately appear to be a dominant factor, it seems likely that the work they reviewed was for traffic signs varying in size by only a small amount; additionally such signs tend not only to be smaller but may generally be less visible than advertising posters since they may not be able to capitalize on the graphic options available for poster design. The signs erected for the purpose of the Cole and Hughes field research ranged in area from 0.07 to 0.38 m^2. Indeed one aim of the present research is to determine if size does have a major role in visibility when varied over the rather larger range such as characterizes the contrast between a bus-stop poster panel and a large wall hoarding (which vary between 2.16 and 37.2 m^2). It will be appreciated that we may well obtain a different outcome given that the *smallest* poster panel used in the present study was more than five times bigger in area than the *largest* disc used in the Cole and Hughes study. Size may well matter for poster panel visibility.

Method

Methodological overview

It will be clear from the introduction that the principal interest of the study centres on whether or not observers of a scene look at any poster or posters it contains. While this could be approached by asking judges independently to assess whether or not people would look at a poster in a scene, a subjective method would not be ideal if an objective technique could be found. In fact this was the case since an eye movement recording facility was available and so this was the method of choice.

The task for the observers was to view a series of still photographs on a computer monitor while their eye movements were recorded. The photographs depicted scenes with roadside settings; most of the scenes contained advertising posters but a proportion of them had no poster in order to avoid the actual purpose of the study being obvious. The observers were asked to view each picture (which was displayed for six seconds) as if they were a driver or passenger of a vehicle in the setting depicted by the photograph.

The instructions used aimed at inducing a viewing style associated with the visual behaviour associated driving or being a passenger (subjects served in only one of these conditions). No reference was made to advertising posters. At the end of the eye movement recording phase the observers were questioned about what they had seen. The instructions served as a kind of cover story for the study; it was important not to convey to the subjects that the research question was whether they looked at advertisements. To refer to poster panels would have invited the subjects to look consciously *for* them as opposed to looking unwittingly *at* them.

The presentation interval of six seconds was based partly on how long a poster panel might be in view for a driver. At an average driving speed of about 10 mph in London (where the photographs were obtained) a distance of 50 metres would be traversed in about 11 seconds. This distance is within the visual reach of most poster sites, but the interval proved to be too long for the presentation of stationary images of typical urban roadside environments; observers had seen enough of these representative but rather dull scenes after a few seconds, the task risked becoming more obviously unrealistic, and a pacier arrangement was clearly needed to sustain interest. It was also necessary to support the cover story, which required the subject to appraise the scene for driving hazards. A much shorter interval of two or three seconds was insufficient for the latter purpose and six seconds was the eventual compromise. This is an aspect of the research that merit further examination in due course.

What was inspected, and for how long, as they inspected each scene was recorded (subjects typically make approximately 12-20 fixations in the viewing interval, their eyes remaining at rest about 80-90 per cent of the time). The scenes varied according to the design constraints described below. Poster content was not controlled, but was assessed on a *post hoc* basis. Photographs were commissioned of sites via a database maintained by NOP.

Subjects typically make 12-20 fixations in the viewing interval, their eyes remaining at rest about 80-90 per cent of the time.

Materials: Scene photography and scene selection

A total of 86 photographic scenes were used: this comprised 76 images containing a "target" (i.e., a poster panel) and 10 decoy (distractor) scenes which had no target. Of the 76 images containing a target, there were respectively 27, 43 and 6 containing a $2m^2$ (6 sheet) panel, a $18m^2$ (48 sheet) panel, and a $37m^2$ (96 sheet) panel respectively.

These images were drawn from a folder of 143 viable candidate images commissioned by Postar for the study: 70 in this folder included at least one 6 sheet panel; 49 included at least one 48 sheet panel; the remaining 24 decoy candidates contained no poster panel.

The photographer for the images was briefed very carefully to supply pictures of scenes containing poster panels of the specified sizes, which were "natural" and representative of scenes that a driver would typically encounter; it was stressed that "marketing shots" of panels would not be acceptable. The distance from viewer to panel was to be 30 metres, which could be verified from screen measurements relative to the calibration images described in the next paragraph; in the event measurements from the scenes as depicted on the screen used the average viewer/camera-to-panel distance was estimated to be 35 metres using screen measures.

Initially two poster locations, a small bus-stop 6 sheet and a wall mounted 48 sheet, were photographed at decreasing ranges in steps of 10 metres from 100 metres to 10 metres, with the camera in a road position estimated to that of a car driver. From these scenes an optimum distance was selected (30m) from which all further photographs would be taken. The images for a measured 10 metres distance were used as screen calibration images.

Next a professional photographer was briefed to take photographs from a similar road-centred and ranged position. The photographer was also instructed to take realistic busy daytime and night-time scenes that a driver would be likely to encounter. The night-time photographs were all taken at dusk so that a reasonable amount of background light was still available, more representative of the scene seen by the eye than a night-time shot. The photographer also required a guard, so that there were no traffic accidents while shots were aligned.

Using a database of location addresses in London, which contained a mixture of billboard sizes and locations (residential, arterial and shopping) some 150 locations were selected and photographed. This process took place in March and April 1995 and contained street scenes of varying weather (and hence light levels) and congestion. Some locations were visited twice: the first was in daylight and shots were taken of the billboard with its internal light both on and off; the second was at dusk/night with the internal light on. The photographer was also asked to take shots of distractor scenes where no billboards appeared, in all three location types and times of day.

The resulting photographs were then sorted by eye and the most representative scenes chosen to fill the experimental design matrix (see Table 1a below).

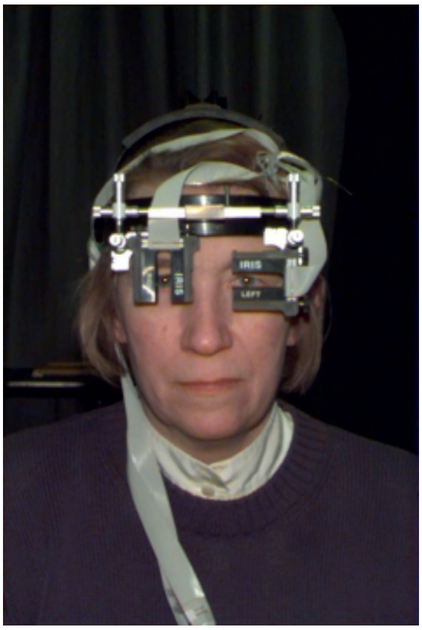
Equipment

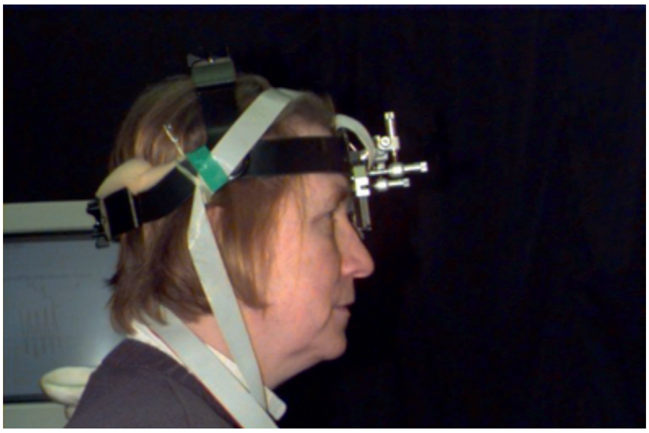
Each subject's eye movements were recorded using a Skalar IRIS eye tracking system interfaced to a Mac Quadra 950, fitted with a National Instruments Lab-NB board for additional input/output operations. This entails the wearing of a headset as shown in Figure 2. A standard monitor was used that supported displays of 768 x 512 pixels. The subject was also supplied with a microswitch to trigger successive trials. Custom-built software was used to control the experiment, data collection and data analysis.

Notwithstanding the fact that computerization of the eye movement record was intrinsic to the Skalar system used, scoring of the data for a single subject was not fully automatic (see Data Analysis section) and took about 45 minutes per subject. A comprehensive description of the visual behaviour was supplied for further analysis, including the coordinates of the screen locations visited by each subject's eyes. It was thereby possible to identify the incidence and duration of fixations on any poster panel contained in a scene.

Despite recent developments in eye-movement recording techniques not all subjects are usable for the eye movement recording process. The basis for eliminating subjects is partly to do with the calibration process (for example, some are unable to comply with the calibration procedure, possibly because they have difficulties in exerting the conscious control over their eye movements that is required for equipment calibration), or because of structural factors, or because they wear spectacles (which impede the recording process). The expected wastage/ rejection rates are about 10-20 per cent. The restriction on wearers of spectacles may be expected to be removed in future studies once the technical platform for eye movement recording is upgraded.

Figure 2: Front and side views of Skalar IRIS eye tracking system





Method
continued

Procedure

There were two research assistants who shared the tasks of meeting the subjects, instructing them, conducting the calibration process, running the experiment, debriefing the subjects, and analysing the data.

On arrival the participants were given a printed instruction sheet briefly outlining the nature of the study and instructing them as to their task. The printed instructions are included as Appendix B, along with the supplementary script used by a research assistant. The participants were invited to ask questions about the task before the session was under way. The session was subsequently divided into two main phases: (1) the eye movement study; and (2) the interview.

(1) Eye-tracking procedure

Each subject's eye movements were recorded using a Skalar IRIS eye tracking system interfaced to a Mac Quadra 950, fitted with a National Instruments Lab-NB board for additional input/output operations. The first step was for a research assistant to calibrate the equipment relative to the individual's eye structures and eye movements. To this end a series of circular targets was presented on the 16 inch (40.6 cm) diameter screen of the Macintosh at random locations sampled from a rectangular matrix spanning the area on the screen on which the poster scenes were to appear. This enabled the data from the sensors mounted on the monitoring frame (head-set) worn by the subject to be interpreted. Each subject sat facing the screen at a distance of 40 cm using a chin-rest mounted on the edge of the table to support the head and to restrict any head movements tending to destroy calibration settings. Initial adjustments of the headset were made to ensure comfortable viewing conditions. Calibration ensued with readjustments (and further calibration) of the frame as necessary. This phase was sometimes protracted as subjects adjusted to the equipment, and the experimental situation. The longest calibration phases were about 15 minutes. Calibration was subsequently checked after every 15 images during the experimental phase.

The 86 poster scenes were presented in a single block, the subject having the option of a rest pause if needed. Each scene was displayed for 6 seconds and the subject's eye movements were recorded during this time. The screen was then blank until a small black square appeared in the centre of the screen as a fixation guide. When ready to proceed the subject pressed a micro-switch key placed on the table in front of the screen and the next scene appeared. Between blocks the subject's comfort and the equipment calibration were checked. At the end of the study the monitoring equipment was removed and the subject went to the interview room for debriefing by a second research assistant.

(2)Interview

The subject was debriefed relative to the purpose of the study immediately after the eye movement phase, so that the interview could focus on the presence of advertising posters in the scenes that had just been viewed. The subject was asked to report anything seen during the drive that directly related to the instructions given at the outset of the study, hence to report information about any hazards they had noticed. The responses typically referred to cars pulling out behind other vehicles, people crossing the road, and so forth. These responses were not analysed since the task was only to ensure that the subject paid due and normal attention to the visual scene.

Audience factors

An important aspect of this study is its examination of the audience factor. This refers to the fundamental distinction between people who are exposed to poster sites as drivers, passengers (and pedestrians, the next target audience), and whose viewing behaviour may be expected to vary. Evidence suggests that passengers are not dramatically different from drivers, however, they are much less visually constrained and may be expected to scan the world outside the vehicle more widely than drivers do (pedestrians are even less visually constrained). Accordingly the study was intended to reflect these likely behavioural differences by giving the subjects instructions that encouraged visual styles to simulate those of interest. Thus "Drivers" were instructed to think of themselves as driving a car, and to begin viewing at a specified viewing point (on the road ahead) and were asked to give a relevant report (at the end of the experiment, for example, how many traffic lights were seen, what hazards were spotted). By contrast, "Passengers" (and "Pedestrians" eventually) have to think of themselves as such, with a suitable "memory task". The task of a driver is relatively straightforward and limited. By contrast a passenger is typically much less constrained in what he or she does or can look at: a passenger may adopt or be assigned the task of navigating, whereas another may assume the role of "back-seat driver", and yet others may simply be there for the ride. We attempted to cover the range of variations by using two versions of instructions. The audience factor varied on a between-subject basis (i.e., each subject served in only one condition). This was reflected in the recruitment process for the study.

Subjects

These were recruited from NOP staff supplemented by participants recruited from the local University of London campus. For the purpose of assessing the audience factor for posters a total of 40 subjects were tested: allocated to a driver group of 17 subjects and to a passenger group of 23 subjects.

Data analysis

The eye movement recordings were analysed by a semi-automated procedure. For each subject's data on a given trial the software displayed a thumbnail sketch and a synchronized record of the subject's horizontal and vertical eye movements. Such recordings typically contain occasional artefacts resulting from eye-blinks and the software included automatic procedures to recognize and ignore these sources of error. However, the recognition algorithm was not perfect and the operator had to intervene manually when an eye-blink artefact was not picked up by the software. This was achieved by using a function key to ignore any affected tract of data. This facility was infrequently required. The visual appearance of an eye-blink is quite evident to an operator, being characterised by a sharp divergence normally in the vertical record, the trace commonly disappearing off the edge of the screen.

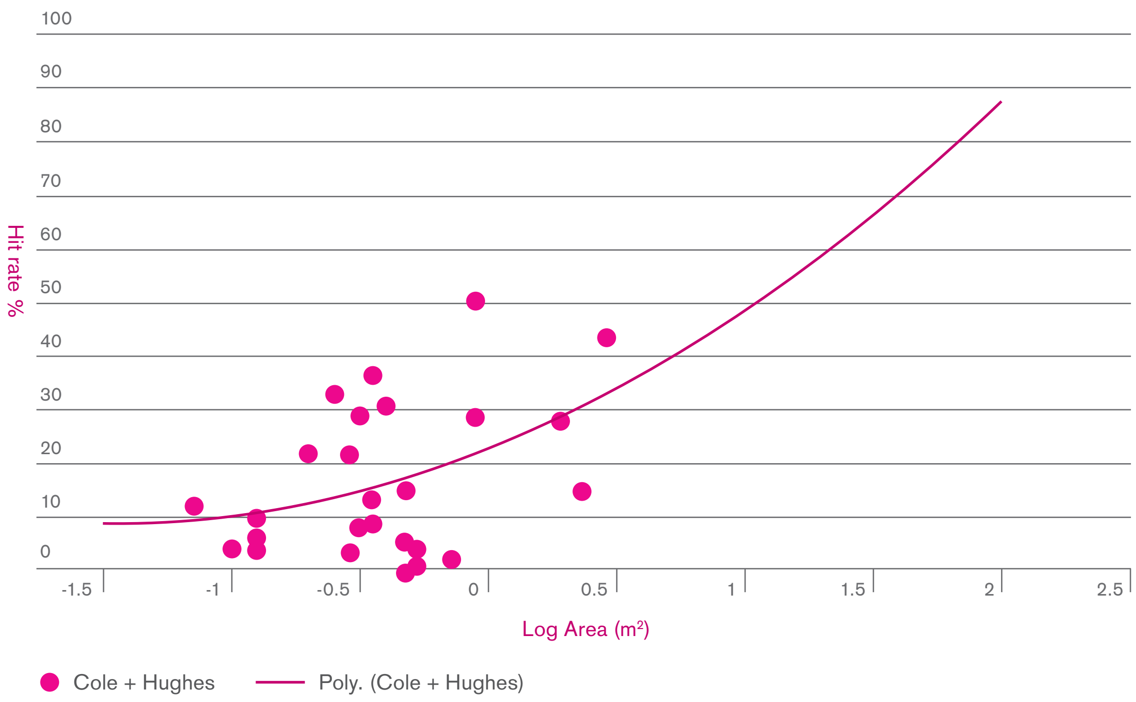
The output of the analysis program was a list of locations fixated and fixation durations in sequence. Each fixation was automatically tagged as a "hit" or a "miss" by reference to a set of target locations, corresponding to the coordinates of the corners of the poster. These coordinates were established separately using a procedure that allowed the operator, by manipulating a cursor on the screen, to note the locations of the corners of the poster panel with the image *in situ* in the host scene.

The analysis reported here is focussed on key aspects of the data that support the development of visibility metrics. However, much more may be extracted from the rich data set potentially available, including evidence of any typical "scanpaths" (regular eye movement patterns between regions of interest in the image) that lend qualitative support to the statistical analyses.

The key "hit rate" measure was obtained by noting whether the poster depicted is fixated. The "latency" (reaction time) of the first fixation (if any) on the poster was also recorded along with the "gaze duration" of such a fixation. The total "gaze duration" of all such fixations and the overall number of hits were also logged.

The response surface for hit rate as a function of eccentricity and size will be estimated, the likely form of which is depicted in Figure 3 (following a section of data drawn from an experiment by Coles and Hughes). The surface is likely to change position vertically depending on various factors; it will probably shift upwards for illuminated panels, reflecting the weight to be attached to this special subset of sites. The moderating effects of clutter on the response surface will also be estimated. The design offers sufficient degrees of freedom for estimating the model[1]. The variability of the data is inevitably unknown at this stage but this will limit how well the model parameters can be estimated.

Figure 3: Hit rate: Attentional conspicuity of road signs (data derived from Cole and Hughes, 1984)



1 Thanks are due to Paul Harris of NOP for assistance in verifying this property of the design.

An overriding consideration is that the demands on the research participants must be realistic.

Reality checks: Some caveats regarding methodology and scale of study

Before describing the details of the research design, it is important to outline some practical considerations that must be taken into account.

(1) Validity of image presentation methods

The images to be used were photographs scanned directly from the negatives to CD-ROM minimizing loss of detail. These were then transferred to a Macintosh Quadra on which they were to be displayed to the subjects as digitised images on a high resolution computer screen. Using a suitable lens (22 mm), an attempt will be made to produce a broader image than is characteristic of 35 mm cameras because the natural visual field width of the human eye is as much as 150° of visual angle. Notwithstanding, the display as seen by an observer will not be as complete as seen in the real world. What drivers/passengers/pedestrians look at could probably only be established by using the most advanced technology, which was not available in the time frame and budget available. For this purpose basic equipment costing between £50k and £60k would be needed, but it is not clear that even this would be able to handle in-car recording of eye movements of real-world scenes.

(2) Ultravision

Some important factors and questions have to be side-lined: thus the question of what weighting factor to apply to Ultravision panels had to await improved technology. For the time being, it was possible only to note the possible pros and cons of such systems (e.g., the possible attentional advantage of movement vs. the loss of information as the panel turns) that would be taken into account by a weighting factor.

(3) Contrast

A second issue that is bound to be highly important but could not be tackled in this study was the effect of visual contrast. It is apparent that poster sites vary in the contrast between the poster itself and its background. This is crudely captured by the notion of "clutter", but for a given class of site, posters will still vary markedly in global contrast relative to the surroundings. It should be noted that the database for OSCAR II was not planned to include contrast information for individual poster sites. The illumination factor exemplifies how contrast can be successfully manipulated and possible even maximised, though non-illuminated sites vary in contrast too.

(4) Distance

The contribution of observer-target distance is clearly of interest and this is another potentially major factor in poster panel visibility. Clearly adding distance as an extra factor in a balanced research design along with panel size, environment and eccentricity is ruled out on practical grounds – the scale of such a study would simply be impossible to achieve. It is also evident that the use of images containing close-up views of poster panels is risky because anyone participating in the study would be likely to infer its purpose and possibly change their viewing behaviour one way or another – deliberately searching for or avoiding advertising material, either of which would be a departure from the normal habitual visual behaviour that needs to be captured. Distance is therefore put to one side as a factor to be explored in a future possible investigation.

(5) Scale of study: Demands on participants

An overriding consideration is that the demands on the research participants must be realistic; they must not be so burdened that they become bored or fatigued. Experience suggests that no more than an hour of testing is the upper limit for a study such as proposed here. Recruiting participants for more than one test session is a possible way of obviating the problem but this has to be a requirement that applies to all participants – posing a considerably greater logistical difficulty than recruitment for a single session – so this was not pursued further.

There are obvious consequences for the amount of data available from a single test session as to be implemented here. The session begins with an eye-position calibration phase of up to 15 minutes; this is followed by a practice phase of about 10 minutes to instruct and familiarize the participant with the procedure; finally there is an eye movement recording phase for the acquisition of data, lasting about 30 minutes. The latter stage also includes rest pauses and calibration checks. With rests and calibration checks lasting about 12 minutes, a total of about 18 minutes remains to be dedicated to data collection. This enables the presentation of just over 100 scenes (of duration 6 seconds each plus a 4 second interval between scenes). In the event a total of 86 scenes was used so that participants were not over-burdened and the procedure could be assured of a smooth passage.

(6) Scale of study: Statistical considerations

An important additional factor in designing the experiment is its internal scale. How, taking statistical factors into account, does the investigator decide on the size of the study? In particular how many images are needed per condition of interest? How many participants should be tested? The ideal is to know a fair amount about the data in advance so that the number of observations taken can be optimised to detect any differences between conditions that are statistically significant. The full formal requirements include a reliable knowledge of certain key properties of the data about to be collected, as might be available from past research or a pilot study. The latter was not feasible for the present research but insights were available from the considerable literature on eye movement research that guided decisions about the scale of the present investigation. Finally the ideal position for the researcher is to know the magnitude of any effects (e.g., percentage differences between panel sizes) that would be important for the client to detect but this was not known at the outset of the study.

(7) Scale of study: Precedents – field survey vs. laboratory experiment

It is probably fair to say that the experience of Postar's user and contributor community would bias expectations regarding sample sizes in the direction of opinion surveys. As a background for the design of research this is bound to contrast with the style of experimental psychological research on which this study is based. There are many methodological options on which experimental psychology relies, and the scale of the resulting research studies varies accordingly. One of its traditions is to follow the lead of psychophysics where important findings may be established by the use of a single observer; for example, in the scientific discussion of light and dark adaption **(see http://webvision.med.utah.edu/book/part-viii-gabac-receptors/light-and-dark-adaptation/)**, several of the results are based on studies on this scale. Of course results do become more precise as sample sizes increase but the gain is generally marginal if the variance in the system measured is sufficiently small. Experiments using eye movement recording are typically of an intermediate size, reflecting the intrinsic variability of the parameters to be estimated. Although many insights have been revealed by the examination of data from single individuals such as the scanning patterns (eye-tracks) evoked in response to pictorial stimuli (Yarbus 1967), the bulk of the eye movement research on reading and the perception of pictures has relied on single- or two-figure sample sizes. Measures that are estimated in such studies (see Rayner 1978) include fixation duration, fixation/saccade frequency and saccade length. The scale of studies such as these provides important guidance for the present study.

Experimental design

The principal independent variables for the study were object (poster) size, eccentricity, clutter and audience. They are fundamental to the proposed visibility model. It should be stressed that the design structure is not fully mirrored in the analysis of results. While an optimal design might incorporate all combinations of the chosen levels of the three factors, this was not possible in the light of the absence in the real world of certain poster sites (e.g., 6 sheet panels at 40° eccentricity).

In practice, key values for size correspond to the industry standards of 6 sheet, 48 sheet and 96 sheet panels; it should be noted that this means that object shape (aspect ratio) and object centre both differ with size. Practically relevant values for eccentricity are 0° (kerbside), up to 15°, 15-45°, 45-75°; these correspond to information logged in the NOP database from the poster interview process. In the present context, "clutter" refers to the ambient environment, known to the industry as arterial vs. residential vs. shopping. It is assumed to covary positively with the amount of visual clutter in the vicinity of the poster site. Ideally a study would be made to develop an independent metric of clutter for this study; meantime environment is used as a surrogate for clutter.

Planned research design and some compromises

A complete factorial study (not including the audience factor) for the combination of size x eccentricity x environment would consist of 3 x 4 x 3 = 36 conditions which might be of theoretical interest but not always practically achievable (for instance, since 6 sheets viewed at their optimum visibility distance are never positioned at 75° offset). The final design structure as planned was determined by the industry's imperatives and the portfolio of billboard options. The aspects to be included in the eventual design are subsumed in one final structure (Table 1a). Note that in the Table, cells that are to be represented by data are shown by the # symbol; blank cells denote that no data would be obtained.

The effect of panel size could be examined by comparing kerbside sites (0° eccentricity) with 6 and 48 sheet posters, but this is subject to the availability of sites. The 48 sheet size was chosen for the purpose of estimating the effect of eccentricity (with values of 0°, 20° and 60° from kerbside into the left visual field). It would also be desirable for a set of images for the 96 sheet size to be obtained for the low clutter (Arterial environment) setting for eccentricities of 0°, 20° and 60°; the data for this set of conditions would be used to test whether extrapolation from the 6 vs. 48 sheet comparison is reasonable.

Another independent variable – not included in the foregoing characterization of the study – that is of major interest is *illumination*. There are panels that are internally illuminated and this may be investigated by obtaining photographs of those sites with and without internal illumination just before dusk, to give a test of the effect on visibility of daytime illumination. The effect of illumination after dusk would also have practical relevance.

Method
continued

Table 1: Proposed study design for visibility variables, combining all factors and factor levels (NI = non-illuminated, I = illuminated)

Panel size	Clutter	Eccentricity 0°			Eccentricity 20°			Eccentricity 60°		
		Day/NI	Day/I	Dusk/NI	Day/NI	Day/I	Dusk/NI	Day/NI	Day/I	Dusk/NI
6 sheet	Shopping	#	#	#						
	Arterial	#	#	#						
	Residential	#	#	#						
48 sheet	Shopping	#			#		#	#		
	Arterial	#			#		#	#		
	Residential	#			#		#	#		
96 sheet	Shopping									
	Arterial	#			#			#		
	Residential									

Summary of research design as delivered

A summary of the design of the study is included as Table 1a (for convenience, the contrast for "dusk" is abbreviated, and does not show the full set of possibilities). The data that are generated from this design enable a considerable majority of the relevant combinations of variables to be sampled. The design is constructed so that the effects of the different factors could be examined by analysis of variance techniques; and so that the "response surface" for visibility can be estimated statistically by multiple regression techniques, to enable the OSCAR II model to be put in place. The gaps in the matrix correspond to conditions that do not exist in the real environment or are very uncommon and may reasonably be discounted. This presents a very challenging prospect for the identification of poster sites to meet these complex criteria. Indeed in a number of respects reality – as indicated by the portfolio of photographs supplied for the study – proved that not all of the effects of the factors of interest could be satisfactorily assessed. Moreover the clutter factor would ideally be underpinned by independent support for its validity in the context of this study. Uncertainty about the design variables and the absence of a full audit of poster sites meant that the study must to some extent be seen as exploratory.

Distractor conditions

To avoid subjects becoming aware of the purpose of the study, and in particular to prevent them from realising that the main interest of the study is in whether they fixate on or near posters, some poster-neutral images will be included in the sequence. There would need to be daylight distractor scenes (decoys) in each Shopping, Arterial and Residential settings, with no posters in view. Further distractors would be required to sample sites at dusk and others in shopping and arterial settings at night but with street lighting.

Results

Preliminaries: final design and raw data preparation

The final experimental design used 86 photographs of street scenes including 10 distractors, presented to 40 subjects (17 drivers and 23 passengers). Their fixations and saccades were recorded and later analysed to remove eye-blinks and other noise, leaving a data set when aggregated and transferred to a spreadsheet of some 62782 rows, each row containing among other information the screen coordinates of that fixation. These fixations were then overlaid by computer onto the target posters in the photographs and a hit or miss recorded; for the purpose of this study any secondary hits were ignored. Moreover for visibility measurement the focus was on the question of whether or not they looked at the target, not on the associated dwell time (which is likely to be affected more by content).

The photographs were then analysed to record the distance from the road centre line to the leading edge, and the x and y coordinates of the corners of each target billboard. From this information each billboard was assigned to an eccentricity band, which increased in steps of 10 degrees eccentricity from 0 to 90 degrees. The width and height and subsequent area of each billboard can then be calculated in pixels as well as the angle subtended at the eye.

Raw data example

Column 1 Subject number (01-52)
Column 2 Picture code (01-86)
Column 3 Fixation number (1,2+)
Column 4 Panel size (6, 48, 96)
Column 5 Day-time vs. night-time (D or N)
Column 6 Hit vs. miss (H or M)
Column 7 Environment (A, R, S) arterial, residential or shopping
Column 8 is redundant
Column 9 internally illuminated / not illuminated (I or NI)
Column 10 Start time of fixation (t1)
Column 11 End time of fixation (t2)
Column 12 Duration of fixation (t2-t1)
Column 13 X-coordinate of fixation (1-768)
Column 14 Y-coordinate of fixation (1-512)

Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Col9	Col10	Col11	Col12	Col13	Col14
05	01	01	006	D	M	R	00	I	0005	0275	0270	0387	0260
05	01	02	006	D	M	R	00	I	0295	2155	1860	0431	0223
05	01	03	006	D	M	R	00	I	2170	2490	0320	0455	0227
05	01	04	006	D	M	R	00	I	2510	2965	0455	0481	0218
05	01	05	006	D	M	R	00	I	2985	3155	0170	0447	0214
05	01	06	006	D	M	R	00	I	3195	3790	0595	0258	0225
05	01	07	006	D	M	R	00	I	3835	4340	0505	0422	0211
05	01	08	006	D	M	R	00	I	4390	5390	1000	0592	0189
05	01	09	006	D	M	R	00	I	5420	5500	0080	0610	0177
05	01	10	006	D	M	R	00	I	5520	5995	0475	0634	0159

Main findings: Effects of panel size, environment and eccentricity

The key factor whose effect on visibility needs to be estimated is panel size, however, this may depend *inter alia* on factors such as road environment (clutter) and eccentricity. The data for panel size are presented first, followed by any interaction effects (with Environment and Eccentricity) that may be of interest.

The visibility of a panel is indexed here by its hit rate, calculated as the proportion of subjects with at least one fixation on the panel, and this is the dependent variable in the following analyses. The tables show mean hit rates, aggregated over subjects and panels for the various panel properties and categories.

The tables also show the standard deviations and number of entries per cell. Data have been omitted for any cell in the table that have less than five entries. One consequence of this is that the individual 96 sheet means are not shown when a factor (i.e., environment or eccentricity) is added. This does not apply to Table 2 which shows mean hit rates as a function of all three panel sizes, with standard deviations and sample sizes).

Figure 4: Mean hit rate as a function of panel size (error bars based on MS error from analysis of variance)

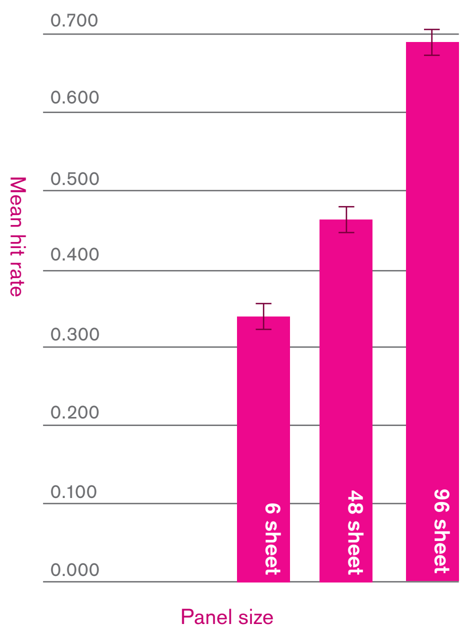


Table 2: Hit rate as a function of panel size

Panel size	6 sheet	48 sheet	96 sheet
Mean hit rate	**0.338**	**0.465**	**0.692**
Standard deviation	0.097	0.139	0.193
Number of panels	27	43	6

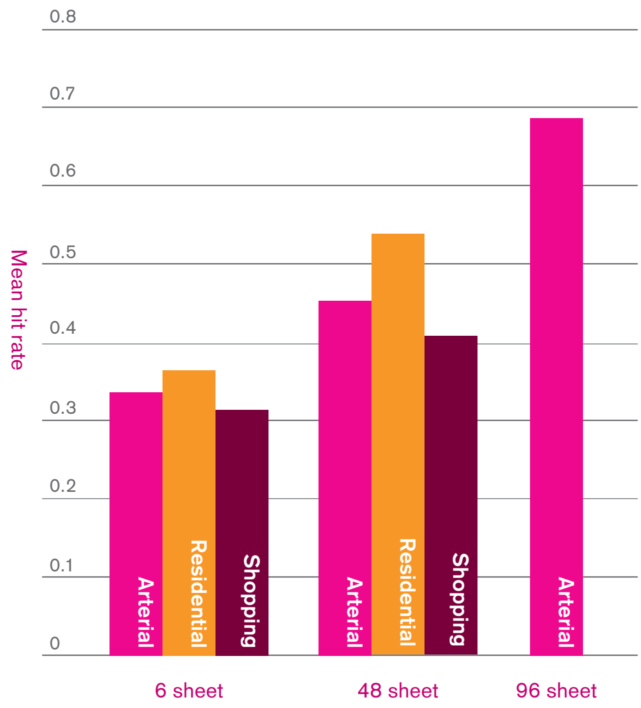
An analysis of variance of hit rate with panel size as the only factor showed that Panel Size is statistically significant ($F(2,73) = 20.10$; MS error = 0.0171; $p<0.001$). Figure 4 depicts the mean hit rates with error bars. Multiple comparisons using the Bonferroni correction indicated that the three condition means were significantly different from each other.

Hit rate as a function of panel size and environment/clutter is examined next. Table 3 presents the mean hit rates along with standard deviations and cell counts (see also Figure 5). The main effects of Panel Size and Environment were both significant according to an analysis of variance with Panel Size and Environment as factors: for Panel Size $F(2,69) = 19.97$; MS error = 0.0160; $p<0.001$), and for Environment $F(2,69) = 3.054$; MS error = 0.0160; $p=0.054$). The interaction of these two factors was not significant ($F(2,69) = 0.540$; MS error = 0.0160; $p = 0.59$). It should be stressed that the design in this respect is unbalanced as a result of the absence of 96 sheets in any but the Arterial road setting. Bonferroni tests revealed that the marginal effect of Environment reflects the lower hit rate scores for Shopping – arguably the most cluttered of the three – compared to Residential and Arterial.

Table 3: Hit rate as a function of Panel Size and Environment

		Panel size			
Environment		**6 sheet**	**48 sheet**	**96 sheet**	**Total**
Arterial	**Mean hit rate**	**0.344**	**0.452**	**0.692**	**0.468**
	Standard deviation	0.106	0.145	0.193	0.188
	Number of panels	9	14	6	29
Residential	**Mean hit rate**	**0.364**	**0.538**	**–**	**0.470**
	Standard deviation	0.108	0.133	–	0.149
	Number of panels	9	14	–	23
Shopping	**Mean hit rate**	**0.306**	**0.408**	**–**	**0.370**
	Standard deviation	0.075	0.116	–	0.113
	Number of panels	9	15	–	24

Figure 5: Hit rate as a function of panel size and environment



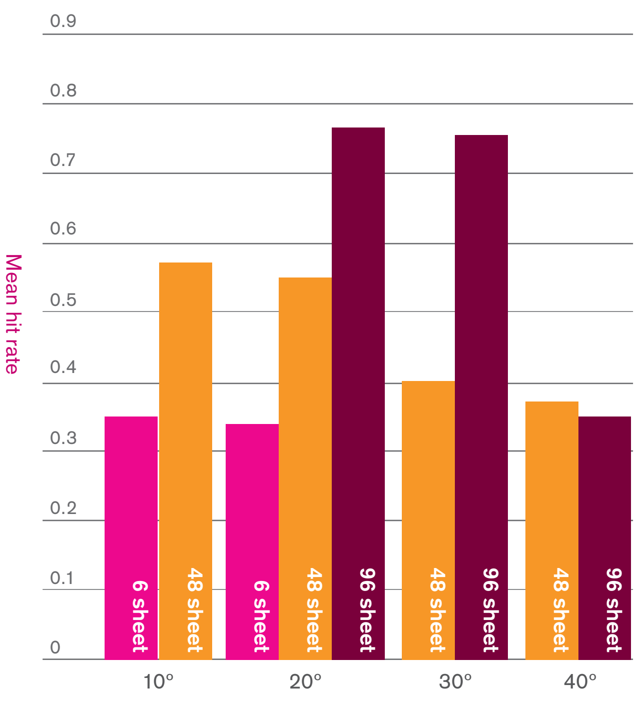
The next factor to consider – eccentricity – is the least tractable of those under review. In principle it would be possible to contrive a study to investigate it in a more controlled fashion than has been achieved in the present study. The challenge of identifying a collection of panels that met the requirements of a complete factorial design seems even greater than that of selecting panels for a full investigation of the effect of environment. In the latter case, locating 96 sheets in Shopping or Residential settings proved to be difficult. Finding 6 sheet panels at anything beyond 30 degrees offset from the edge of the road would be even harder, maybe impossible. The fact is that the research has to be based on what is there. It transpired that in the photographs of panel sites, the panels were located at eccentricities up to about 60 degrees. Eccentricities were grouped into 10 degree bands for the purpose of analysis and the numbers of panels in each band for each panel size are reported in Table 4 (and are depicted in Figure 6). The table also shows mean hit rates with the corresponding standard deviations, as in the previous tables.

Table 4: Hit rate as a function of panel size and eccentricity

Panel size	Eccentricity band	10°	20°	30°	40°
6 sheet	**Mean hit rate**	**0.345**	**0.331**	**–**	**–**
	Standard deviation	0.095	0.103	–	–
	Number of panels	14	13		
48 sheet	**Mean hit rate**	**0.575**	**0.549**	**0.405**	**0.375**
	Standard deviation	–	0.129	0.115	0.177
	Number of panels	1	17	23	2
96 sheet	**Mean hit rate**	**–**	**0.763**	**0.758**	**0.350**
	Standard deviation	–	0.053	0.146	–
	Number of panels	–	2	3	1
Total	**Overall mean hit rate**	**0.360**	**0.473**	**0.446**	**0.367**
	Overall standard deviation	0.109	0.173	0.163	0.126
	Number of panels	15	32	26	3

A superficial look at the data in Table 4, focussing on the row reporting the data pooled over panel sizes conveys the impression of an inverted U-shaped curve, hit rate declining with eccentricity after an initial rise between 10° and 20°. This is not consistent with the impression that may be gained by reading Table 4 one panel size at a time, which hints at a decrease in hit rate with eccentricity for each panel size; the means for each panel size decline numerically as angle band increases. This is a liberal account considering that the decrement in the 6 sheet is minimal. It should nevertheless be noted that the inverted U-shape, with a sharp initial increase in the *combined* hit rate, is largely down to the differential composition of the means for the first two eccentricity bands – the mean of 0.360 for 10° is weighted substantially by low-scoring 6 sheet panels, whereas the mean of 0.473 for 20° reflects a more even weighting by 6 sheet panels and high-scoring 48 sheet panels.

Figure 6: Hit rate as a function of panel size and eccentricity



Plainly because of the uneven distribution of panel sizes it is difficult to draw conclusions about the effect of eccentricity. There were sufficient data for partial analysis of the panel size x eccentricity band matrix, using the best populated cells in the matrix, which throw light on the effect of eccentricity. Thus two separate analyses of variance of 6 sheet scores for 10° and 20° and 48 sheet scores for 20° and 30° were conducted. For the 6 sheet analysis the decline between 10° and 20° bands was – in percentage terms – a mere 1.4% and not surprisingly this effect was not statistically significant ($F(1,25) = 0.134$; MS error = 0.010; $p = 0.72$). By contrast the percentage decline of 14% between 20° and 30° bands for the 48 sheets was highly significant ($F(1,38) = 13.66$; MS error = 0.015; $p = 0.001$).

Viewpoint differences: Driver vs. passenger

It seems likely that the viewpoint of the driver – and therefore an active participant – in contrast with that of a passenger – and thus a more passive occupant of the vehicle – will materially affect what elements of a scene are looked at. The contribution of the role assigned to the subject was therefore assessed statistically by adding the factor of Role (Driver vs. Passenger) to the analysis of variance for Panel Size. For the analysis the three panel sizes and represented by different pictures but the same pictures are viewed by both drivers and passengers. Hence Panel Size is considered to be a between-items (or pictures) factor and Role is treated as a within-items factor.

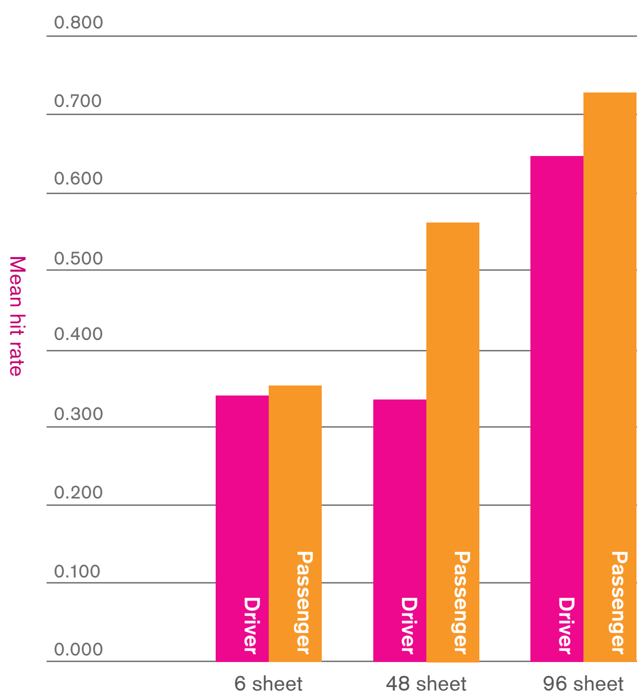
Both Panel Size and Role were significant (Panel Size $F(2,73) = 17.77$; MS error = 0.0358; $p<0.001$); Role $F(1,73) = 25.65$; MS error = 0.0088; $p<0.001$). The interaction of these two factors was also highly significant ($F(2,73) = 24.91$; MS error = 0.0088; $p<0.001$).

Table 5: Hit rate as a function of panel size for drivers and passengers

Viewpoint	Measure	Panel size		
		6 sheet	48 sheet	96 sheet
Driver	Mean hit rate	**0.333**	**0.330**	**0.647**
	Standard deviation	0.140	0.174	0.244
Passenger	Mean hit rate	**0.341**	**0.564**	**0.725**
	Standard deviation	0.087	0.142	0.167
	No. of scenes	27	43	6

The source of the effect of role and the interaction effect is seen in Table 5 and in Figure 7, the graphic version of the mean hit rates. Passenger hit rates are appreciably greater for passengers than for drivers but only for the two large panel sizes.

Figure 7: Hit rate as a function of panel size: Drivers vs. passengers



This leaves three analyses to consider. First, an analysis using role and environment as factors with a potential to interact showed nothing of note (for Role x Environment F(2,73) = 1.643; MS error = 0.0142; p = 0.20). Second, there are two analyses mirroring the previous analyses with band as the factor of principal interest, but adding role with the potential to interact. The analyses treat 6 sheets (10 vs. 20) and 48 sheets (20 vs. 30) separately. As before, they are conducted separately because of the design imbalance that arises from the availability of these two panel sizes in the real world. For the 6 sheet results, none of the main or interaction effects were significant: Role (F<1), Band (F<1) and Role x Band (F(1,25) = 3.00; MS error = 0.0063; p = 0.096). For the 48 sheet data, Band was significant (F(1, 37) = 12.64; MS error = 0.0104; p < 0.001), as was Role (F(1, 37) = 99.47; MS error = 0.0104; p < 0.001). The Role x Band interaction was not significant (F<1). Focussing on the interactive contributions of Role, it is evident that this viewpoint factor does not significantly modify the effect of eccentricity (Band). Any other effects have been considered earlier.

Growth of hit rate over time

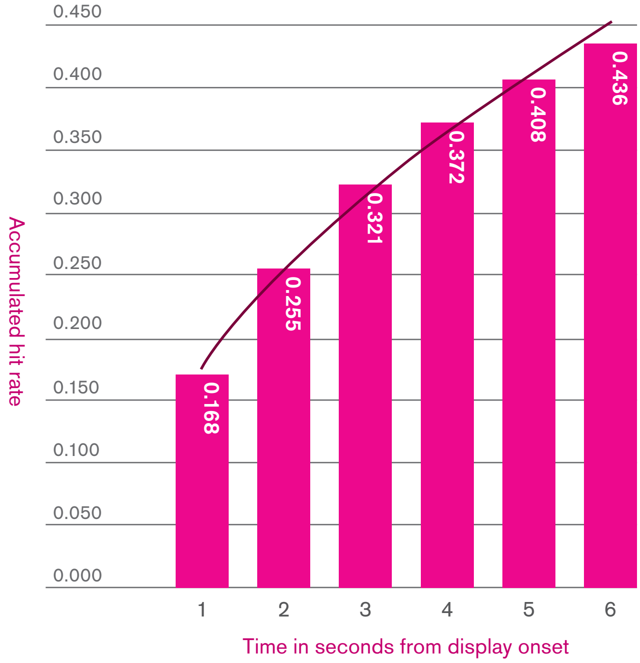
It is evident growth of hit rate is loaded towards the beginning of the display interval. This is underlined by data on how hits are distributed through the interval. Altogether the 40 subjects scored 2924 hits; of this total 1324 were first hits and 1600 were contingent hits (second or later). The distribution of these hits over the 6-second display interval is of interest. The lower or first quartile (25% percentile) of the "first hits" was achieved after 0.56 seconds, the median (50% point) was reached after 1.50 seconds, and the upper or third quartile (75% percentile) was reached after 3.08 seconds. Of the 1600 contingent hits 718 were second hits, and the three quartiles were reached at 1.23, 2.58 and 4.40 seconds.

It was reported above that hit rates for passengers were higher than for drivers and this is reinforced by the differential hit rate growth rates demonstrated by the two groups: the first hit quartiles for passengers were at 0.42, 1.16 and 2.90 seconds while for drivers they were at 0.95, 2.18 and 3.68 seconds. It seems that inspection of the scenes – at least for salient objects such a poster panels – is weighted towards the first two or three seconds of the display. This is reinforced by data on the mean latency of the first hits – that is, when on average did first hits occur[2]. The aggregate mean latency, pooled over subjects and pictures was 1.96 seconds; for drivers and passengers respectively the mean latency was 2.44 and 1.71 seconds.

These findings suggest that visual exploration of the scenes is not evenly distributed throughout the display interval, but is strongly loaded in the direction of the start of the interval. The literature on eye movements suggests that observers of a fresh scene tend to have exhausted their visual exploration of the scene quite early on (Mackworth and Morandi, 1967), so it is of interest that this phenomenon is demonstrated in the present study. More evidence concerning the distribution of fixations on target objects during the presentation interval is presented in the remainder of this section.

It is of course obvious that hit rate will grow as the display interval proceeds, reaching its final value after the 6 seconds used for presenting the photographs. The build-up of hits is shown in Figure 8, in which the hit rate is shown at the end of successive seconds. It can be seen that hit rate, which reaches a high of 45.4% after 6 seconds is well on the way (almost 20% by the end of the first second, with successively smaller increments after this. These successive increments from the first second on are 8.7, 6.6, 5.1, 3.6 and 2.8. The initial surge in hits has ended by the mid-point of the display interval.

Figure 8: Hit rate accumulated as a function of exposure time (seconds) pooled over all conditions; hit rate is also shown as a power function estimated from the data



2 It will be noted that median latencies are rather lower than the means; as different types of average they reflect different features of the distributions of data (in this case, the distributions show positive skew).

Results
continued

Hit rate is well expressed as a power function of time (hit rate = 0.173 T 0.537) as shown by the overlaid curve (r=0.994). By extrapolation on this function, the hit rate reaches 100% after 26 seconds.

If this is extended – with extreme caution because of the sample sizes involved – to the data for panel size, very similar profiles are obtained (see Figure 9). The equations of the curves are not shown to avoid too much clutter but they all are associated with very high correlations (0.988 and above); by extrapolation hit rate for 6 sheet panels reaches 100% after 53 seconds, 48 sheets just after 21 seconds and 96 sheets after 18 seconds. A power function was chosen as it provided a very good fit, but other functions – including logarithmic and polynomial functions – would also serve well. The choice between function types is best made with a supportable theoretical rationale, as is considered in a later report.

A possibly clearer way of viewing how hit rate progresses is to decompose the accumulated hit rate data for each second of exposure as shown in Figure 8; with the scores pooled across panel sizes. This is repeated to show the equivalent pictures for each panel size (Figure 9). The data points are simply the increments second by second derived from the data in the previous figures. Hit rate is highest in the first second and there is a substantial drop between the first two seconds. A power function is overlaid on the data, again fitting the data well. Nonetheless other more complex functions need to be explored.

Figure 9: Hit rate accumulated as a function of exposure time (seconds) for each panel size (pooled over environments) and estimated power functions for hit rate vs. time

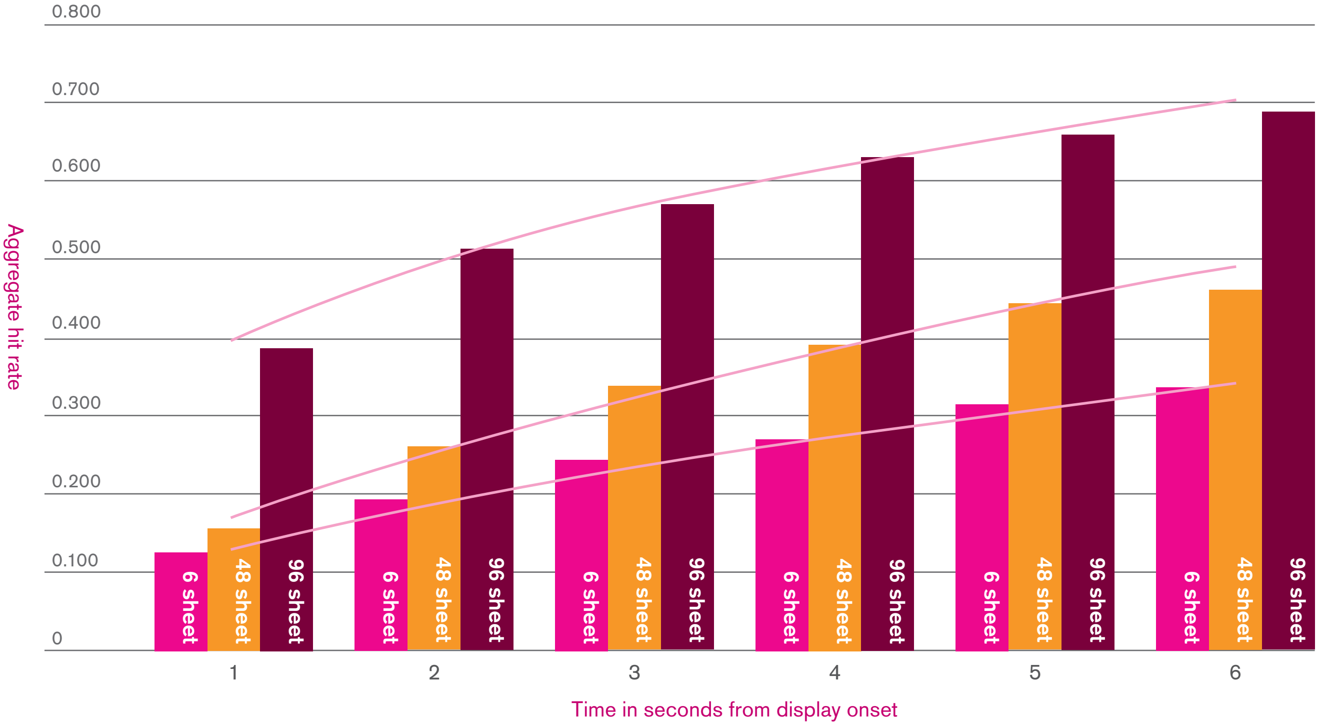


Figure 10: Hit rate changes per second of exposure, aggregated over panel sizes; hit rate increments are also shown as a power function estimated from the data

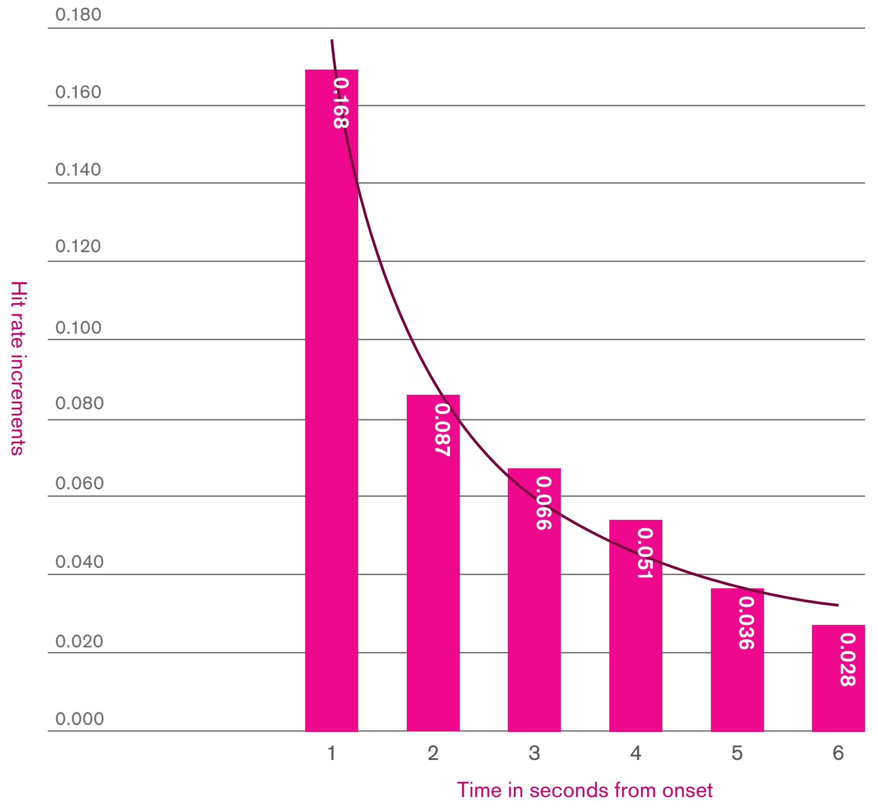
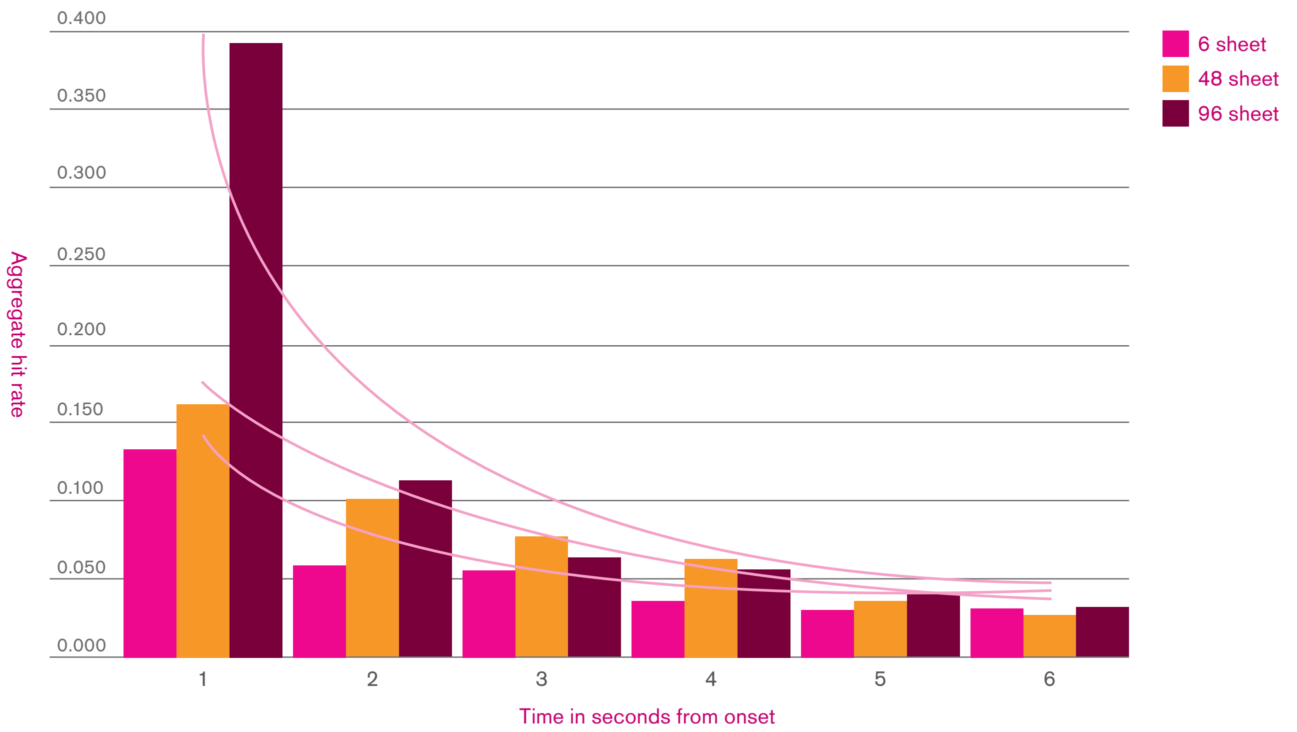


Figure 11: Hit rate per second for the three panel sizes: curves shown are for power functions of hit rate vs. time



Curve fitting and preamble to visibility modelling

As a basis for our own investigations we extracted and transposed some of the results of the Cole and Hughes study (1984) which showed hit rate as a function of target area. On these data we have overlaid a regression curve (in Figure 12), which could be used to predict the likely results from our own investigation, with the potential to guide some preliminary visibility modelling. Actual data from our experiment are added in the figure, demonstrating the level of agreement between various aspects of the predicted and experimental hit rate results. Of course this is somewhat tenuous given the contrast in the research domains (real-world driving and scenes vs. laboratory role-playing and photographic scenes).

Figure 12: Conspicuity of road signs (data from Cole and Hughes, 1984): with the addition of data from present visibility study, and showing estimates to illustrate predicted hit rates for key panel sizes

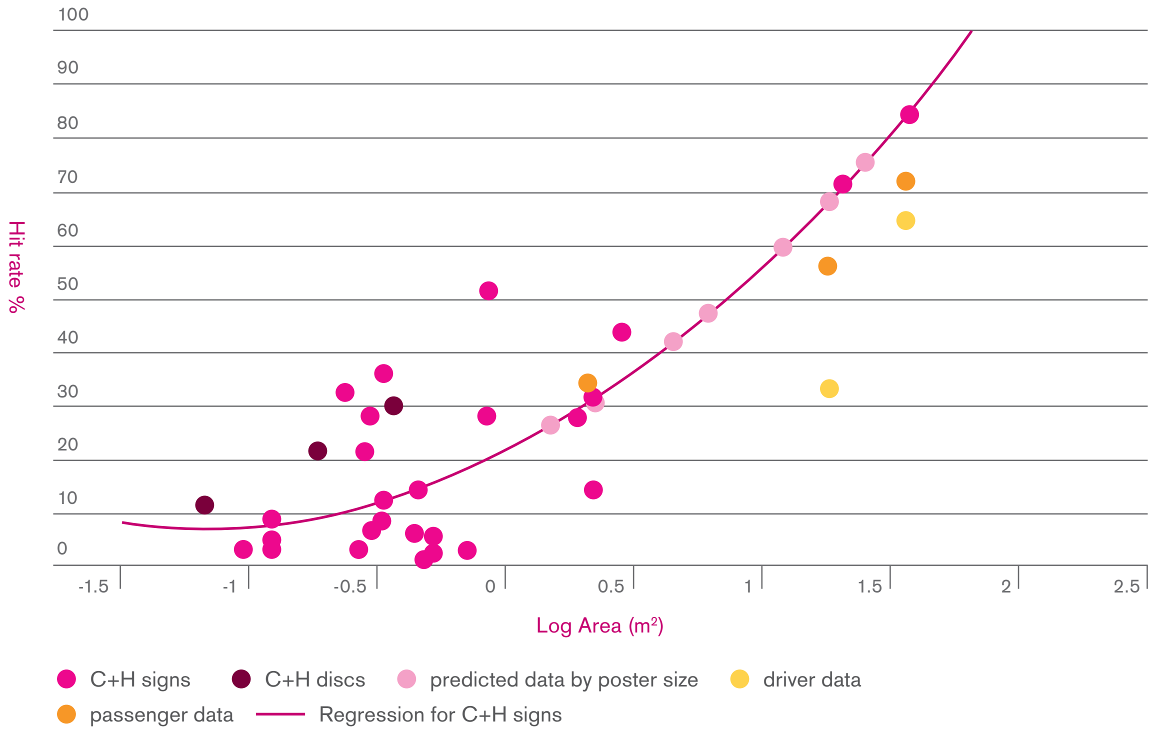


Figure 12 shows a non-linear regression curve (second order polynomial) based on Cole and Hughes data set. This is an approach to the hit rate by size calculation that could be incorporated in a visibility model. As an indication of the effect of panel size, Figure 12 also shows the hit rate of some selected poster sizes when the regression formula is applied. The fit between the present data and the regression curve is certainly sufficient to warrant further exploration of the hit rate vs. panel size relation. It is moreover clear from the earlier analyses of the effects of panel size on hit rate that size does matter, and to a greater extent than indicated by Cole and Hughes (1984) study, in which a considerably more limited range of sizes was employed.

To further illustrate the fit between predicted and recorded hit rates the data points plotted in this Figure include the results for drivers and passengers from the present study. The deviations are greatest for the two larger panel sizes. Of course there are manifold reasons for such an outcome (mostly arising from the methods used and including the contrast between laboratory and field settings for the research), and the measures of hit rate employed, however, the exercise points to the task to be faced in developing a visibility model, and the technical domain within which solutions may be sought.

It is evident that a model based on the account proposed by Cole and Hughes (1984) may provide a satisfactory visual fit for the data from the study. Notwithstanding it should be remembered that the data involve an aggregation over factors such as road type/environmental clutter (as indexed by the distinctions between arterial, residential and shopping settings) and it is possible that such factors may require different model parameters, or even separate models. Moreover the hit rate predicted from the polynomial model may need to be derived for the viewing conditions used; in particular this includes the viewing interval used, and the depiction of the scene at a single distance. The reality is that a given scene is only seen for a fraction of this time, and what is seen – by a driver, for example – is a continuously unfolding view. A poster panel in the scene is approached, perhaps directly, maybe at an angle; and its image changes accordingly. It maps out a visual trajectory and does so in real time as determined by the motion of the vehicle (and thus of the viewer). The problem of accounting for what may be characterised as an aggregation of successive scene glimpses is addressed in a later report.

The fit between the present data and the regression curve is certainly sufficient to warrant further exploration of the hit rate vs. panel size relation.

Conclusions and recommendations

The objective of this research has been to provide the empirical foundations for a model of the visibility of roadside poster panels. This is to be used to adjust raw measures of the relevant "audience". The empirical yield of the study is complex and extensive. The data was obtained by using eye movement recording equipment to monitor the visual behaviour of observers acting as car drivers or passengers. These observers viewed a collection of scenes, most of which featured a poster panel. For a given scene/panel the measure of visibility derived from the raw data was the incidence of fixations (hit rate) on the target object – that is, the panel (if any) contained in the scene.

The method of eye movement recording was implemented successfully. Many aspects of the data could be reasonably judged as having "made sense", as indicated by the various empirical regularities reported and summarised below. The instructions to those taking part to view the scenes as driver or passenger appear to have been successfully conveyed, as shown by the hit rate findings comparing the two roles.

The key findings of the study were that:

- Hit rate tended to increase as panel size increased
- Hit rate was lower for panels in Shopping scenes than those in Arterial or Residential scenes
- Hit rate tended to decrease as angular offset from the centre of vision increased but this was mostly due to a sharper decline beyond 20° offset
- Hit rates achieved by Passengers tended to be greater than those achieved by Drivers but this was largely attributable to an advantage for larger panels
- Hit rate accumulated most rapidly in the first second or two of the viewing interval, and this was less marked for Drivers (who must attach the greatest priority to the road directly ahead) than Passengers (who inevitably are less restricted in what they can inspect and when, and can therefore visually explore with more freedom).

The regularities in the data, relative to the key factors of panel size, offset and environment, as well as the evidence pertaining to the viewing roles adopted by the observers, provide a promising starting-point for the development of a visibility model for poster panels. There is nevertheless much work to be done for such a model to be confidently put in place.

References

Barber, P.J. and Legge, D. (1976). *Perception and information*. London: Methuen.

Cole, B.L., and Hughes, P.K. (1984). A field trial of attention and search conspicuity. *Human Factors,* **26**, 299-313.

Cole, B.L., and Hughes, P.K. (1990). Drivers don't search: they just notice. In D.Brogan (Ed.), *Visual Search*. London: Taylor & Francis, 407-417.

Cole, B.L., and Jenkins, S.E. (1982). Conspicuity of traffic control devices. *Australian Road Research*, **12**, 223-238.

Cole, B.L., and Jenkins, S.E. (1984). The effect of variability of background elements on the conspicuity of objects. *Vision Research*, **24**, 261-270.

Engel (1971) Visual conspicuity, directed attention and retinal locus. *Vision Research*, **11**, 563-576.

Johansson, G., and Rumar, K. (1966). Drivers and road signs: A preliminary investigation of the capacity of car drivers to get information from road signs. *Ergonomics*, **9**, 57-62

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

Mackworth, N.H. and Morandi, A.J. (l967). The gaze selects informative details within pictures. *Perception and Psychophysics*, **2**, 547 552.

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

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin*, **85**, 618–660

Sanders, A. (1963). The selective process in the functional visual field. Soesterberg. The Netherlands : Institute for perception, RVO-TNO, 65,66.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, **74**.

Unema, P. and Rötting, M. (1990). Differences in eye movements and mental workload between experienced and inexperienced motor vehicle drivers. In D. Brogan (Ed.), *Visual Search*. London: Taylor & Francis, pp. 193-202.

Yarbus, A. L. (1967). *Eye movements and vision*, New York: Plenum.

Acknowledgements

We are most grateful to Dr Tim Hodgson for providing the software for the empirical study; which included the experimental control procedure, as well as the eye-tracking processes and data analysis. Thanks are also due to Brian Aviss for technical assistance, and the Department of Psychology at Birkbeck College, University of London for hosting the research.

Appendix A: Why "visibility"

An early requirements-analysis for the study revealed an imperfect consensus as to what the research was intended to measure; in preliminary discussions there was an impossibly wide range of views as to terminology – and meaning. Terms were canvassed ranging from the detectability of a panel to its impact. Other concepts were mooted including noticeability, visibility, and conspicuity. It was agreed that something of the sort was what should be measured, if possible.

The concept of impact seems better suited as capturing what an observer recalls from a poster, what is remembered of it, or what action is taken as a consequence of seeing it, and this hinges on the poster not the panel – the execution or content, as opposed to the panel itself independent of content. The other variant – detectability – is according to the Encarta Dictionary the capacity to "notice or discover the existence of something". But this implies a deliberateness or consciousness of action, and it was agreed that the aim was to ascertain if a panel was attended to. In short, was the panel seen? Did the person with a possible view of the panel actually look at it? Note that the question posed is not whether the panel could – with an effort to do so – be seen. The further assumption is that once attention has been triggered it may continue to be directed, or to be redirected, towards the object responsible.

Two of the remaining options, "noticeability" and "visibility", may not really be distinguishable, and both potentially raise the question of intention. As to noticeability the difference (in intention) is reflected in the contrast between the questions "is the panel noticed?" vs. "can the panel be noticed?" The same contrast seems less easy to apply if the term "visibility" is adopted; moreover, it has the singular merit of emphasising that the visual system at the core of the matter. In addition there are precedents in the relevant research literature (see below) in which an object's visibility indexed by measures of whether or not the viewer looks at the object. For completeness we should mention "conspicuity", another strong contender, which figures prominently in vision research and is cited a number of times in this paper. Though it has the benefit of neutrality it has the drawback that it is absent from certain dictionaries and thesauruses. What tipped the scales in favour of visibility was its clearer link to vision.

Appendix B: Instructions

These consisted of a sheet of printed instructions (in italics) and supplementary instructions given orally by the research assistant. The oral instructions are shown with quotations marks.

i: Experiment on visual behaviour of drivers

This experiment is to discover what people in cars look at in photographic scenes. We are doing this by recording eye movements using an infrared eye movement tracker. This sends out a very faint harmless infrared signal (less than any domestic coal-fire or radiator) that is reflected off the dark edge of the eyeball and is detected by a tiny sensor on the tracker. This is a completely safe and standard procedure. To set up the equipment before the pictures are presented, we have to go through a short calibration sequence. This is so that the computer can interpret the signals that the tracker picks up from your eyes, which are unique to each individual. The calibration sequence is a series of circles on the screen which you have to look at in a certain pre-set order. This is a rather boring but essential part of the experiment that may take as long as 15 minutes. It takes time to be precise. Once that's done the experiment will begin.

In the experiment, while your eye movements are being recorded, you will be shown a series of about 80 road scenes. On the screen is a typical picture.

Think of yourself as a driver, looking down the road ahead, steering the car and watching out for hazards in the normal way. You will have a few seconds to look at each picture, and we'd like you to view the scene each time, just as you would when driving. For each scene, make up your mind as soon as you can as to whether the road conditions are safe or not, and then carry on viewing whatever catches your eye, as if driving along that road. There's no right or wrong answer, just your judgement, but you'll probably find that almost all of the situations are really quite safe. As soon as the picture goes off, signal your decision by pressing the right-hand key for "safe" or the left-hand key for "not safe". The computer will display the next picture after a few more seconds, so be ready. Every so often there will be a short break to check the settings of the eye tracker.

"Do you understand what's involved? Do you have any questions? ....
Then let's begin by calibrating the equipment."

Calibration sequence here

"We are now ready to do the experiment, and we next have a few slides to get you used to what's involved. When each picture comes on look first at where you would be focussing if you were driving. Then carry on, judging if it's a safe situation, and looking at anything else that you might view, allowing for the fact that you will need to drive safely."

Practice slides here

If all goes to plan .....
"That's good. OK, let's do the next batch of slides."

End/Debriefing

"That's all. Thank you very much for taking part. Your results will be put together with those of about 50 others to draw up a final picture of what people look at under these circumstances, and particularly whether they look at the poster signs in the displays. My colleague will deal with the payment."

Appendix B: Instructions continued

ii: Experiment on visual behaviour of passengers in cars

This experiment is to discover what people in cars look at in photographic scenes. We are doing this by recording eye movements using an infrared eye movement tracker. This sends out a very faint harmless infrared signal (less than any domestic coal-fire or radiator) that is reflected off the dark edge of the eyeball and is detected by a tiny sensor on the tracker. This is a completely safe and standard procedure. To set up the equipment before the pictures are presented, we have to go through a short calibration sequence. This is so that the computer can interpret the signals that the tracker picks up from your eyes, which are unique to each individual. The calibration sequence is a series of circles on the screen which you have to look at in a certain pre-set order. This is a rather boring but essential part of the experiment that may take as long as 15 minutes. It takes time to be precise. Once that's done the experiment will begin.

In the experiment, while your eye movements are being recorded, you will be shown a series of about 80 road scenes. On the screen is a typical picture.

(Version 1): *Think of yourself as a passenger in a car, looking down the road ahead trying to keep a look out for road signs so you can help give the driver instructions on where to go. You will have a few seconds to look at each picture, and we'd like you to view the scene each time just as you would when being driven in a car. For each scene, make up your mind as soon as you can whether there is road sign information to view more closely, and then carry on viewing whatever catches your eye, as if being driven along that road. You'll probably find that almost all of the situations contain no road sign information.*

(Version 2): *Think of yourself as a passenger in a car, looking at the scene ahead. You will have a few seconds to look at each picture, and we'd like you to view the scene each time just as you would when being driven in a car. For each scene, make up your mind as soon as you can whether it is a familiar place, and then carry on viewing whatever catches your eye, as if being driven along that road. You'll probably find that most if not all of the situations are unfamiliar.*

The computer will display the next picture after a few more seconds, so be ready. Every so often there will be a short break to check the settings of the eye tracker.

"Do you understand what's involved? Do you have any questions? ....
Then let's begin by calibrating the equipment."

Calibration sequence here

"We are now ready to do the experiment, and we next have a few slides to get you used to what's involved. When each picture comes on look first at where you would be focussing if you were being driven as a passenger in a car, probably down the road ahead. Then carry on, judging if there's a road sign to look at, and looking at anything else that you might view as a passenger."

Practice slides here

If all goes to plan .....
"That's good. OK, let's do the next batch of slides."

End/Debriefing

"That's all. Thank you very much for taking part. Your results will be put together with those of about 50 others to draw up a final picture of what people look at under these circumstances, and particularly whether they look at the poster signs in the displays. My colleague will deal with the payment."

Appendix C: Visibility studies undertaken to date

Driver visibility study (1995-1996): OSCAR 2 measuring visibility hit rates of roadside panels, using infra-red eye-tracking methodology. Introduced the basic concept of visibility hit rates for poster panels. Modelled visibility in terms of panel size, eccentricity (offset from road) and distance. Respondents: drivers and passengers.

Maximum visibility study (1996-1997): assessing the furthest distance at which a panel can be seen with full concentration on the panel, using psychophysical methods.

Pedestrian visibility study (1998-1999): measuring visibility hit rates for poster panels in roadside and pedestrian environments, using infra-red eye-tracking methodology. Respondents: pedestrians.

Nottingham driver attention study (2000-2001): establishing how drivers' & passengers' attention is distributed down the road ahead – using real-world in-car eye camera technology. Respondents: drivers and passengers.

"Inclusivity" pilot (2002): comparing a set of active search methods as alternatives to passive eye-tracking methods (for speed, convenience and portability).

Wave 1 (aka Travel Wave) (2003-2004): using an active search method selected on the basis of the "Inclusivity" pilot to estimate hit rates for panels from transport media (buses, tube, rail, taxi). Respondents: pedestrians.

Wave 2 (aka Retail Wave) (2003-2004): using the active search method to estimate hit rates for panels in retail environments (supermarket car-parks, malls, pedestrian shopping precincts, petrol stations, telephone kiosk). Respondents: pedestrians.

Video analysis of driver eye behaviour (2004-2005): using video analysis of gaze data from Nottingham driver attention study to assess hit rates on roadside panels and buses. Respondents: drivers and passengers.

Pedestrian visual behaviour: walking speed and head-up study (2005): specifying key aspects of walking for use in pedestrian visibility modelling via literature searches and observational data.

Wave 3 (2006): using the active search method to provide supplementary data on panel hit rates in key transport environments (buses and tube). Respondents: pedestrians.

Wave 4 (2007-2008): using a passive eye-tracking method to estimate panel hit rates in key transport and retail environments, with contemporary roadside panels, providing an up-to-date database across environments with new eye camera technology. Respondents: drivers and pedestrians.

Wave 5 (2008): a passive eye-tracking method to update estimates of panel hit rates for telephone kiosks and taxis. Respondents: drivers and pedestrians.

Dynamic Imagery Research Phase 1 (2008-2009): Pilot study to explore technology for presenting moving images (scrolling displays) while recording eye movements. Respondents: unclassified.

Dynamic Imagery Research Phase 2 (2009): Investigation of effect of dynamic images (scrolling poster panels and bus panels) on hit rates, using a stationary view of the scene. Respondents: pedestrians.

Dynamic Imagery Research Phase 3 (2009-2010): Investigation of effect of dynamic imagery (scrolling and digital poster panels, and bus panels) on hit rates, using a dynamic view of the scene. Respondents: drivers and pedestrians.

Visibility of poster panels seen through bus and train windows (2010): using a passive eye-tracking method to estimate panel visibility when viewing through a bus or train window. Respondents: pedestrians and public transport passengers.

route
outdoor media analytics