Developing Predictive Equations to Model the Visual Demand of In-Vehicle Touchscreen HMIs

David R. Large[1], Gary Burnett[1], Elizabeth Crundall[1], Editha van Loon[1], Ayse L. Eren[1], Lee Skrypchuk[2]

[1]*Human Factors Research Group, University of Nottingham, Nottingham. UK.*
{david.r.large; gary.burnett}@nottingham.ac.uk
lizzie.crundall @gmail.com
{editha.vanloon; ayse.eren}@nottingham.ac.uk
[2]*Jaguar Land Rover Research, International Digital Laboratory, University of Warwick, Coventry, UK,*
lskrypch@jaguarlandrover.com

Running Head: developing predictive equations to model visual demand

ABSTRACT

Touchscreen HMIs are commonly employed as the primary control interface and touch-point of vehicles. However, there has been very little theoretical work to model the demand associated with such devices in the automotive domain. Instead, touchscreen HMIs intended for deployment within vehicles tend to undergo time-consuming and expensive empirical testing and user trials, typically requiring fully-functioning prototypes, test rigs and extensive experimental protocols. While such testing is invaluable and must remain within the normal design/development cycle, there are clear benefits, both fiscal and practical, to the theoretical modelling of human performance. We describe the development of a preliminary model of human performance that makes *a priori* predictions of the visual demand (total glance time, number of glances and mean glance duration) elicited by in-vehicle touchscreen HMI designs, when used concurrently with driving. The model incorporates information theoretic components based on Hick-Hyman Law decision/search time and Fitts' Law pointing time, and considers anticipation afforded by structuring and repeated exposure to an interface. Encouraging validation results, obtained by applying the model to a real-world prototype touchscreen HMI, suggest that it may provide an effective design and evaluation tool, capable of making valuable predictions regarding the limits of visual demand/performance associated with in-vehicle HMIs, much earlier in the design cycle than traditional design evaluation techniques. Further validation work is required to explore the behaviour associated with more complex tasks requiring multiple screen interactions, as well as other HMI design elements and interaction techniques. Results are discussed in the context of facilitating the design of in-vehicle touchscreen HMI to minimise visual demand.

CONTENTS

1. INTRODUCTION

There has been significant interest in the theoretical prediction of human behaviour and performance, and a substantial corpus of literature exists. Applied to the field of human-computer interaction (HCI), this work has traditionally focussed on technology and interfaces that act as the only or primary focus for a user's attention (e.g. menu selection using desktop computers), with the aim of predicting interaction time or user performance. However, given that the world is now suffused with computers and technology, users may be required to interact with an interface while concurrently executing a more critical primary task (e.g. while driving). Consequently, there is increasing interest in the prediction of human behaviour and performance in situations which place multiple demands on users' attention. Existing work in this domain has centred on the development and integration of individual cognitive frameworks to make *a priori* predictions concerning behaviour for both primary and secondary tasks (e.g. Salvucci, 2001), and typically necessitates a detailed and thorough understanding of the task, the user and the system, and the interactions between them (Harvey et al., 2011). However, early models often failed to capture individual differences in the desire or ability of users to engage in a secondary task, while maintaining satisfactory primary task performance, and are equally absent of any efficiencies achieved through repeated exposure to an interface (i.e. learning effects). While significant advancements have been made (e.g. Pettitt, et al., 2007), there is general consensus that the prediction of human performance and behaviour in a divided-attention context remains a challenge, with existing 'combined' models often requiring substantial reworking following integration (see: John et al., 2004). In contrast, our work utilises a more rudimentary approach, applying underlying information theory from basic principles to predict secondary task performance.

1.1. Background: Information Theory

A common foundation for modelling behaviour is Information Theory (Shannon and Weaver, 1949), which views humans as information processors. Adaptations of Information Theory were first applied to the then fledgling field of HCI during the 1980s by Card et al. (1983), who articulated two information theoretic models as guiding principles to enhance technology and interface design and usability, notably Fitts' Law and the Hick-Hyman Law (Fitts, 1954; Hick, 1952; Hyman, 1953).

Fitts' Law concerns the prediction of movement time necessary to acquire a visual target (typically using a pointing device) and is predicated on the fact that human performance is limited primarily by the capacity of the human motor system, as determined by the visual and proprioceptive feedback that permits an individual to monitor their own movement and activity. The Hick–Hyman Law (Hick, 1952; Hyman, 1953) compliments Fitts' Law by modelling the relationship between information load and choice-reaction time, i.e. the time taken to determine which target/item to acquire before moving towards it – interfaces with more options have higher information content, as do unlikely events.

Both Fitts' Law and the Hick-Hyman Law are thus highly applicable within the fields of HCI and Human Factors. Fitts' Law, in particular, has been used extensively within the HCI community, for example, as a theoretical framework for computer input device evaluation (Card et al., 1978; MacKenzie, 1992), a tool for optimising new interfaces (Bi et al., 2012; Lewis et al., 1999) and as a method to understand or predict performance when selecting items from menus/interfaces (Cockburn et al., 2012; Cockburn et al., 2007), and is formally recognised in the evaluation of pointing devices (ISO, 2007). However, much of this work overlooks the time taken to choose and locate the target, or fails to consider adaptations to visual search strategies, and associated changes in search-time, as users become more familiar with interfaces. Investigations have also tended to be conducted in a sedentary context, where the interface is the only or primary focus for the user's attention.

1.2. Combining Fitts' Law and Hick-Hyman Law

Given their common root in information theory and obvious practical application, there is an inherent logic and attraction in combining elements of Fitts' and Hick-Hyman models to create a more expansive prediction of human behaviour than each model offers alone. Indeed, since their inception, several attempts have been made to fuse the two models (e.g. Beggs et al., 1972; Soukoreff and MacKenzie, 2004; Hoffmann and Lim, 1997). However, these have achieved limited success, attributed, at least in part, to attempts to incorporate elements that were not amenable to theoretical analysis or intended to be modelled by the laws. Moreover, whereas Fitts' law is immediately applicable to highly familiar and recognisable tasks, such as selecting an on-screen icon or typing on a keyboard, and thus captures human performance that is kinaesthetic and related to dexterity, Hick-Hyman incorporates degrees of unpredictability in stimuli and considers tasks that contain a cognitive element (Seow, 2005). At a theoretical level, the successful application of Hick-Hyman is thus dependent on first codifying different events and then determining their probabilities, in order to calculate their information content or entropy (Landauer & Nachbar, 1985).

Nevertheless, a more recent application by Cockburn et al. (2007) successfully combined elements of Fitts' and Hick-Hyman laws to predict the static task time associated with different menu designs. The Cockburn et al. (2007) model predicts that the time to select an item, T, comprises a Hick-Hyman decision/search time component (T_{dst}) and a Fitts' Law pointing time (T_{pt}):

$$T = T_{dst} + T_{pt} \tag{1}$$

Pointing time, T_{pt}, is derived using a Fitts' Law relationship. To determine item decision/search time, Cockburn et al. (2007) recognise that when users first encounter a menu, they are required to visually search for the target regardless of the layout/design.

However, as users become more familiar with the menu design, benefits may be realised based on their spatial location memory, provided that the placement of items remains predictable and stable. Thus, the model calculates decision/search time by interpolating between a linear visual search-time component and a logarithmic Hick-Hyman decision time component. Cockburn et al. (2007) also model user's expertise, *e*, (with values notionally ranging from 0, 'complete novice', to 1, 'complete expert') to reflect the fact that as user's familiarity increases, their visual search time tends towards zero and the Hick-Hyman decision time component then dominates. Users' experience is determined by the number of previous trials (selections) of the item and the learnability, L, of the interface. Cockburn et al. (2007) nominally assign values for L from 0 to 1, with 1 representing an entirely learnable interface, interpreted as 'items that do not change location or position'. The model therefore recognises that users can never reach 'expert performance' with some interfaces because of poor learnability within the design. It is noteworthy that calculating experience in this manner fails to recognise the number of items in a menu; in practice, it is expected that users will require more trials to become expert with menus containing more items.

The final Cockburn et al. (2007) model thus predicts that the time to select an item (*T*), can be determined by combining the following Hick-Hyman and Fitts' elements (please refer to original Cockburn et al. (2007) paper for further clarification of terms and derivation):

$$T = (1 - e_i)(b_{vs}.N + a_{vs}) + (b_{hh}.\log_2 N + a_{hh}) + \left(a_{pt} + b_{pt}.\log_2 \frac{D}{W}\right) \qquad (2)$$

Results of validation studies conducted by Cockburn et al. (2007) indicated that their predictions of static task time were generally very accurate – within 2% of empirically collected data. However, a limitation of the Cockburn et al. (2007) model – at least from our own research perspective – is that it only applies to the prediction of static task time in a sedentary context.

1.3. Applying Cockburn et al.'s Approach in an Automotive Context

Human-machine interfaces (HMIs) are increasingly prevalent within modern automobiles. From an automotive design perspective, the HMI is a critical customer-facing attribute that represents the touch-point of the vehicle. Automotive manufacturers therefore often employ enticing and aesthetically pleasurable interactive designs, often comprising a touchscreen embedded within the centre console of vehicles, to enhance the driving experience. Indeed, touchscreen interfaces offer a beguiling design solution, typically attracting the most positive opinions from drivers, but they can also be more effective when undertaking common tasks (e.g. menu selection), compared to other devices, such as rotary controllers and touchpads, even while driving (Burnett et al., 2011). Furthermore, the omnipresence of touchscreens in everyday society (i.e. within non-automotive domains) means that such devices are likely to be familiar and ‘intuitive’ to use.

In contrast to a ‘desktop’ sedentary context, however, automotive users are also encumbered by a conflicting primary task of driving. The collocation of visually-enticing HMIs and driving naturally raises concerns. Touchscreen HMIs (in particular) inherently demand some visual attention, due in part to designers’ slavish adherence to skeuomorphic interface elements, even in the automotive domain, to reflect previously physical buttons. Screen layouts and target elements may therefore be visually captivating, and interactions often rely on strong visual cues in lieu of tactile prompts. Users may therefore be forced to visually sample the interface in order to locate and acquire on-screen elements and confirm activation. Vehicle interiors may also be littered with portable, aftermarket devices, such as nomadic GPS-enabled navigation devices and smartphones that require visually demanding touch-based interaction. Evidently, interacting with a poorly designed or visually-enticing HMI, whether it adorns an OEM (original equipment manufacturer) or aftermarket device, has the potential to divert drivers’ attention away from the road scene, and there is little dispute that increased ‘eyes-off-road’ time elevates the risk to drivers, and indeed other vulnerable road users, and can cause deleterious effects on driving performance and vehicle control (NHTSA, 2013).

Consequently, several standardised test protocols have been developed (e.g. ISO, 2014; NHTSA, 2013) that aim to discourage the introduction of highly demanding devices in vehicles. These measure visual demand empirically utilising experimental techniques such as driving simulation and eye-tracking, in which test participants interact with the new device or HMI while driving in a simulator, or visual occlusion, which assesses the visual/manual demand induced by in-vehicle devices or tasks by regulating drivers’ visual activity using shuttered glasses. However, such testing can be expensive and time-consuming to conduct, often requiring extensive user trials involving large numbers of test participants and fully functioning prototype systems (at least in so far as the functionality under examination). While empirical testing is invaluable, and it is imperative that it remains within the design and development cycle, there are clear benefits (both fiscal and practical) to be gained by the theoretical prediction of task time and visual demand associated with new HMI designs intended for in-vehicle placement.

Whereas Cockburn et al. (2007) applied Fitts’ Law and Hick-Hyman Law to predict static task time, the primary concern associated with secondary task engagement within the automotive domain is visual demand – notably recognised by several, widely-cited international driving standards/guidelines (e.g. ISO, 2014; NHTSA, 2013). Vision provides the primary source of information available to drivers and there is a demonstrable link between glance behaviour and safe driving. For example, naturalistic driving studies have shown that the risk of a crash or near-crash event increases significantly as eyes-off-road time increases above 2.0 seconds (Klauer et al., 2006; NHTSA, 2013). Considering the visual demand required to extract salient information from (and interact with) an in-vehicle device or interface is thus considered a direct indicator of the suitability of the interface for in-vehicle deployment (ISO, 2014; NHTSA, 2013).

Secondary task visual demand is typically measured using three key metrics: total glance time (TGT), mean glance duration (MGD) and number of glances (NG). Common definitions for these terms are provided by International driving standards/guidelines. For example, ISO 15007 part 1 (2014) defines TGT as the "summation of all glance durations to an area of interest (or set of related areas of interest) during a condition task, subtask or sub-subtask". MGD is defined as the "mean duration of all glance durations to an area of interest (or set of related areas of interest) during a condition task, subtask or sub-subtask", and NG is the "count of glances to an area of interest (or set of related areas of interest) during a condition task, subtask or sub-subtask" (ISO, 2014).

1.4. Overview of Research

Inspired by the work of Cockburn et al. (2007), we conjectured that a similar approach to that employed by Cockburn et al. (2007) could be applied in an automotive domain. Our research therefore draws upon the influential work by Cockburn et al. (2007) as motivation, but with the notable difference that our model has been applied to single target acquisition using finger-touch input on an in-vehicle touchscreen HMI while driving. We also extend the approach to consider *anticipation* afforded by the structuring of interfaces, and the learning effects of repeated exposure. Moreover, rather than task-time *per se*, we aim to build a model to predict the visual demand of in-vehicle HMIs. To date, there is no evidence of successful attempts to combine Fitts' and Hick-Hyman Laws to model the visual demand of touchscreen HMI interactions in a driving context. The paper thus addresses the following research questions:

1. Do Fitts' Law and Hickman-Hyman Law apply in a driving context?

2. Can the Cockburn at al. (2007) approach be applied to visual behaviour?

In building a predictive model of the visual demand of in-vehicle touchscreen HMIs, we aim to provide both an evaluation and design tool, allowing stakeholders to consider the visual demand of a larger number of HMIs or menu designs, intended for in-vehicle placement, much earlier within the design cycle and without the pre-requisite investment in costly implementation or extensive user trials.

The paper incorporates three studies, which were conducted independently. Section 2 ('General Method') describes the shared methodology applicable to all studies. Section 3 ('Data Collection and Analysis') presents further details for each study and describes the ongoing development of the model.

Section 3.1 describes the derivation of an initial *pointing* component for individual target items. In section 3.2, study two considers the effect of flanking targets on pointing efficiency, thereby revising the *pointing* component to account for arrays of multiple items. The results from study two are also used to derive a *decision/search* component (presented in Section 3.3). In addition during study two, we explore the effect of *structuring* – presenting target items within structured and unstructured arrays. Finally, in Section 4, we present study three, in which the proposed models are validated by comparing predictions with actual performance using a fully functioning prototype in-vehicle touchscreen HMI. Results are presented, analysed and discussed throughout each section.

2. GENERAL METHOD

2.1. Approach

In line with Cockburn et al. (2007), we consider target item selection as comprising *decision/search* and *pointing* components. Cockburn et al. (2007) were concerned with predicting the time to select an item using static menus. Assuming similar logic, we conjecture that the glances associated with selecting an item comprises glances associated with both *deciding/searching (ds)* and *pointing (pt)*:

$$TGT = TGT_{ds} + TGT_{pt} \tag{3}$$

$$NG = NG_{ds} + NG_{pt} \tag{4}$$

For MGD, we assume that this can be obtained from TGT and NG data, in line with International Standards definitions (ISO, 2014), i.e.:

$$MGD = \frac{TGT}{NG} \tag{5}$$

Finally, in order to isolate *decision/searching* and *pointing* behaviour, we initially assume that these activities can be modelled separately from one another. To ensure that participants' behaviour could be segregated in this manner, we asked participants to locate their hand on the steering wheel at a designated location during each study, and remove this only when they began pointing. Therefore, for the purpose of analysis, we assume that *pointing* behaviour began at the time that participants' hand left the steering wheel and anything prior to this (following the presentation of a new stimulus) was deemed to be associated with *deciding/searching*.

2.2. Apparatus and Design

All testing took place using a medium-fidelity, fixed-based driving simulator based at the University of Nottingham. The driving simulator comprised the front half of a 2001 right-hand drive Honda Civic car positioned within a curved screen affording a 270° viewing angle. A bespoke driving scenario was created using STISIM (v2) software to replicate a generic three-lane UK motorway, and projected onto the screen using three overhead projectors (Figure 1). The scenario was based on that specified by the NHTSA eye glance testing using a driving simulator (EGDS) protocol (NHTSA, 2013), but included additional features/infrastructure (e.g. additional road users, geotypical terrain, lateral curvature, UK central 'Armco' barriers and road markings), with the aim of creating a more engaging and ecologically-valid environment. This scenario is well-established, having been employed successfully during numerous similar research studies conducted at the University of Nottingham (see: Large, et al., 2015).

(Figure 1 about here)

The car-following, dual-task paradigm was employed throughout testing: drivers were required to undertake the secondary task (typically, locating and/or selecting a target on the touchscreen) while performing the primary driving task of following a lead car travelling at 70 mph. This approach is typically employed in driver distraction research/testing to control primary task workload, and is specified as part of a standardised experimental protocol within International driving standards (e.g. NHTSA, 2013).

For all studies, participants were self-selecting volunteers, comprising experienced and regular drivers, who responded to advertisements posted online and around the University campus; further details are provided with each study. All participants were reimbursed with shopping vouchers as compensation for their time and provided written informed consent before taking part.

Stimuli were displayed on an HP EliteBook 2740p tablet computer that was located in representative locations within the centre console of the car. During the first study, the location of the tablet computer was alternated between two different positions – notionally referred to as 'upper' and 'lower' – within the centre console of the simulator. In all other testing, the tablet computer was located in the 'upper' position – a more common location for an in-vehicle, centre console display. All testing assimilated data from multiple stimulus-response iterations to ensure that the derived relationships were well-founded and robust. SensoMotoric Instruments (SMI) Eye Tracking Glasses (ETG) were used to collect binocular gaze data at thirty frames-per-second throughout testing.

3. DATA COLLECTION AND ANALYSIS

3.1. Pointing – Single Targets

Twelve people took part in study one: 6 male, 6 female. Participants' mean age was 33, with ages ranging from 21 to 53 years. All participants held a valid driving licence and were experienced and active drivers (mean time with licence, 14 years; range, 1 to 30 years; mean current annual mileage, 6000).

Participants undertook two driving sessions. During one of these, the touchscreen was located in the upper position; for the other, it was in the lower position (Figure 2 shows the tablet computer in the lower position). Participants experienced the touchscreen in both locations, with the order of touchscreen location (upper-lower or lower-upper) counterbalanced between participants. Each driving session lasted approximately fifteen minutes.

At the start of the testing phase of each session, participants were instructed to locate their left hand at a predefined position on the steering wheel marked with white tape (see Figure 2) approximating to the 10 o'clock position on an analogue clock face. This ensured that the Fitts' metric, *D* (distance to target), could be measured for each target location and remained consistent between participants. All testing occurred on straight segments of the motorway scenario. Thus, the location of the tape on the steering wheel remained in the same position in 3-dimensional space for all interactions. The time at which drivers removed their hand from the steering wheel also acted as the demarcation between *deciding/searching* and *pointing* behaviour during the subsequent studies.

(Figure 2 about here)

During each trial, a single square target item appeared on the touchscreen (Figure 2), accompanied by an audible tone to inform participants of the presence a new stimulus. Targets varied in size (6, 12, 18, 24mm) and location, based on existing in-vehicle HMI guidelines. Participants were instructed to touch ('point at') the target as promptly as possible (while maintaining safe driving), and then return their hand to the steering wheel. After touching the target, it disappeared from the screen and, after a short delay, a new target appeared. The order of presentation of target locations and dimensions was randomised between participants.

During study one, the independent variables were: distance to target (*D*) (measured from the white tape on the steering wheel to the geometrical centre of the target) and width of target item (*W*). Dependent variables were: response (pointing) time and visual demand – number of glances (NG), total glance time (TGT) and mean glance duration

(MGD). For the purpose of analysis, we assume that no search time is associated with single target acquisition.

(Figure 3 about here)

Pointing behaviour is obtained from the traditional Fitts' Law relationship ($a + b.\log_2 \frac{D}{W}$), where *D* and *W* represent the amplitude of movement (distance to target) and target width, respectively, and *a* and *b* are empirically-derived model parameters that reflect the efficiency of the pointing system. Observing the ration of distance to target to width of target (i.e. $\frac{D}{W}$) plotted against pointing *TGT* (Figure 3), it is evident that a strong relationship exists. Using regression analysis, the following prediction for TGT_{pt} can thus be derived:

$$TGT_{pt} = 0.26 + 0.11 \log_2 \frac{D}{W} \qquad (R^2 = 0.77) \tag{6}$$

Given that *pointing* is achieved within a single glance (moreover, this is an assumption of our analysis approach), NG_{pt} is assumed to equal 1.0 and MGD_{pt} can thus be determined using the same relationship as TGT_{pt}.

Recognising that touchscreen HMIs are unlikely to contain only one target (in any context), a second study was conducted using larger target arrays to explore the effect on pointing efficiency of additional, flanking targets.

3.2. Pointing – Multiple Target Items

Sixteen people took part in study two: 12 male, 4 female. Mean age was 25.3. All participants held a valid driving licence and were experienced and active drivers (mean time with licence, 7.3 years; range, 1 to 27 years; mean current annual mileage, 5100). Participants undertook seven drives, each lasting approximately 5 minutes. During each drive, participants were required to find and select a single target, located amongst an array of similar items (Figure 4). Target size was consistent throughout the study, in line with existing in-vehicle HMI guidelines, and was the median size used during study 1 (i.e. 15mm).

(Figure 4 about here)

To add a 'decision/search' element, targets comprised white squares containing two letters. Participants were presented with a pre-recorded auditory cue – a target word, spoken aloud – and were required to locate and select the on-screen element containing the first two letters of that word as quickly as possible (example screenshots are provided in Figure 4). Although this search-and-select task may be considered artificial in an automotive context, the approach was chosen to ensure that the effect of structuring could be explored – in this case, alphabetical. Furthermore, by selecting only those words (between 6 and 12 letters in length) that conformed to regular UK English phonetic pronunciation and spelling – meaning in particular that any acoustic ambiguity (e.g. homophones) and/or pop-out effects were avoided – and by presenting the target word immediately before each search task began, the cognitive load associated with remembering the word and matching its first two letters was minimised. An alternative 'ecologically-valid' approach in which automotive icons were utilised would likely have invited too much variability in cognitive load, with participants also required to interpret the meaning of different icons prior to selection.

Target arrays varied in size from one to 36, but targets were always adjacent, equally spaced and grouped in squares, i.e. 1x1, 2x2, 3x3 and 6x6, affording 1, 4, 9 or 36 targets (Figure 4). Targets were presented as either alphabetically-structured arrays, thereby encouraging anticipation, or unstructured arrays. Each of the seven drives constituted a different array size and structure, and these factors remained consistent throughout the drive (i.e. the seven drives corresponded with the seven layouts presented in Figure 4).

Participants experienced all structured array conditions sequentially (i.e. N=1, 4, 9 and 36), followed by all unstructured conditions (or vice versa to avoid order/learning effects), with the order of array size (N) presentation within each condition (structured/unstructured) randomised. During each drive (and therefore each configuration), participants were required to locate and select 36 targets, with targets appearing in each of the possible locations. The layout of the unstructured arrays remained consistent within each of these drives in order to investigate the experience effects of repeated exposure to an unstructured display.

During study two, the independent variables were: anticipation (structured versus unstructured) and array size (number of targets) (*N*) (1, 4, 9, 36). Dependent variables were: secondary task/response time, comprising search time and pointing time, and visual demand – number of glances (NG), total glance time (TGT) and mean glance duration (MGD). Analysis is predicated on the assumption that *decision/searching* and *pointing* activities occur independently and in series and the data were segregated on this basis. As before, it is assumed that pointing began when participants' hands left the steering wheel, and participants were instructed as such during testing; the time that this occurred was obtained using frame-by-frame video analysis. We initially consider the pointing behaviour associated with single target acquisition from an array of multiple square items presented on the touchscreen (from N=1 to 36).

The observed behaviour during the second study suggests that, when only a single item was presented on the screen, pointing behaviour (TGT) conformed with the predictions made by the Fitts' relationships derived during study one (Equation 6), i.e. TGT_{pt} = 6.0 seconds. In the presence of additional, flanking items, however, TGT_{pt} exceeded the predictions from the first study (Figure 5), suggesting that pointing efficiency was degraded in the presence of additional targets, thereby elevating visual demand.

(Figure 5 about here)

The additional visual attention demanded by the presence of multiple targets (highlighted by the lighter shaded areas in Figure 5) was derived empirically as: $0.045 \log_2 N$ ($R^2 = 0.98$). This was concatenated with Equation 6 to enhance our model of pointing behaviour:

$$TGT_{pt} = \left(0.26 + 0.11 \log_2 \frac{D}{W}\right) + (0.045\ \log_2 N) \tag{7}$$

Again, we assume that *pointing* requires only one glance, so NG_{pt} remains equal to one and MGD_{pt} can effectively be determined using the same expression as TGT_{pt}.

3.3. Decision/Search – Uniform Array

In line with Cockburn et al. (2007), we assume that for the visual metrics under investigation, decision/search behaviour can be determined by interpolating behaviour between a linear visual search component (TGT_{vs}) and a logarithmic Hick-Hyman decision component (TGT_{hh}). Cockburn et al. (2007) conjectured that the degradation of the visual search component is determined by users' experience with the particular interface/item under investigation (*e*), with values of *e* range from 0 (representing a complete novice) to 1 (complete expert) (i.e. as user's familiarity increases, the linear search component tends towards zero, and the logarithmic Hick-Hyman decision time component dominates) (Equation 8).

$$TGT_{ds} = (1 - e).TGT_{vs} + TGT_{hh} \tag{8}$$

To model users' expertise, Cockburn et al. (2007) consider the number of previous trials (selections) of the item, *t*, and the learnability, *L*, of the interface (Equation 9).

$$e = L\ \times \left(1 - \frac{1}{t}\right) \tag{9}$$

where values for *L* range from 0 to 1, with 1 representing an entirely learnable interface – described by Cockburn et al. (2007) as one where items do not change location or position – and zero, by inference, a menu that is continually changing and thus impossible to learn.

In contrast to Cockburn et al. (2007), we believe that in a driving context, the ability to *anticipate* the location of the desired target, and *resume* searching after attention has been directed elsewhere (i.e. towards the driving task), are better predictors of visual performance/behaviour than learnability *per se*. While the concept of anticipation may be closely related to learnability, in so far as if an interface is quick and easy to *learn*, then it should also be possible to *anticipate* the location of a desired target (and quickly resume searching) even after limited exposure, we feel that the concept of learnability, as presented by Cockburn et al. (2007), has limited practical application in a driving context. Moreover, while we understand that an *entirely learnable* interface may be assigned a value of L=1, locating other designs elsewhere on a continuum from zero to one appears rather arbitrary and Cockburn et al. (2007) provide limited guidance to support this. Therefore, in our model, we account for 'learnability' through the affordance of anticipation, and have therefore evaluated interfaces that are either *structured (st)* or *unstructured (un)* during the development of our predictive model.

Furthermore, it is expected that users will require more trials to achieve expert status while using menus containing larger numbers of items, particularly in situations of divided attention. Consequently, we would expect the visual demand associated with *searching* (TGT_{vs}) to persist longer in the presence of additional target items and a driving context.

To account for these factors, we used mathematical modelling to define a new visual search/experience scale factor, d_{vs}, applicable to situations of dual task/divided attention (Equation 10). This replaces the expressions for *experience* and *learnability* proposed by Cockburn et al. (2007). Moreover, the new scale factor removes the need to arbitrarily assign numerical values to these constructs.

$$d_{vs} = \frac{\log_2 N}{\log_2(N+t)} \tag{10}$$

$$TGT_{ds} = d_{vs}.TGT_{vs} + TGT_{hh} \tag{11}$$

Determining experience/expertise in this manner also reflects the fact that it is more difficult to anticipate the location of targets (or 'learn' the interface) for larger target array sizes, i.e. as *N* increases, the degradation in the search term is retarded proportionally to the number of items presented. Furthermore, the new scale factor reflects the fact that when N=1, there is no search component (an assumption we also made during study one).

Again, analysis is predicated on the assumption that *decision/searching* and *pointing* activities occur independently and in series and the data were segregated on this basis. We assume that *decision/search* ended when participants' hands left the steering wheel to point, as instructed. We also assume that, at t=1, all participants were 'inexperienced' and at t=36, all participants were 'expert'. While these values may appear arbitrary, the assumption is supported by empirical data which suggests that users' performance levelled out between these extremes.

Visual Search Component

In line with Cockburn et al. (2007), we assume that when a user is inexperienced, they are unable to anticipate item location. Using empirical data, based on participants' initial exposure to the interface (i.e. t=1), we observe a linear relationship between TGT and the total number of items, for both structured and unstructured interfaces/arrays (Figure 6). A similar linear relationship is observed for NG (Figure 7).

(Figure 6 about here)

(Figure 7 about here)

Using these relationships, we can derive the following equations:

$$TGT_{vs\_st} = 0.029N + 0.44 \qquad (R^2 = 0.97)\ (@t = 1) \tag{12}$$

$$TGT_{vs\_un} = 0.10N - 0.028 \qquad (R^2 = 0.99)\ (@t = 1) \tag{13}$$

$$NG_{vs\_st} = 0.021N + 1.04 \qquad (R^2 = 0.98)\ (@t = 1) \tag{14}$$

$$NG_{vs\_un} = 0.044N + 0.81 \qquad (R^2 = 0.94)\ (@t = 1) \tag{15}$$

Hick Hyman 'Decision' Component

As users become more experienced, they are able to anticipate the location of the item based on spatial memory. Using empirical data, based on participants' final exposure (i.e. t=36) – at which point they are deemed to be expert performers – we observe a logarithmic Hick-Hyman relationship between TGT and the total number of items, for structured interfaces/arrays (Figure 8).

(Figure 8 about here)

In contrast, the data obtained from unstructured interfaces suggest that the relationship between glance duration and number of items is linear, suggesting that no such anticipation (or learning) is possible for unstructured interfaces (Figure 8), as might be expected.

Using these relationships, we can derive the following equations:

$$TGT_{hh\_st} = 0.069.\log_2 N + 0.094 \qquad (R^2 = 0.84)(@\ t = 36) \tag{16}$$

$$TGT_{hh\_un} = 0.049N - 0.091 \qquad (R^2 = 0.998)(@\ t = 36) \tag{17}$$

Relationships for NG can be derived in a similar fashion and provide the following:

$$NG_{hh-st} = 1 \qquad (@\ t = 36) \tag{18}$$

$$NG_{hh\_un} = 0.0071N + 0.96 \qquad (R^2 = 0.98)(@\ t = 36) \tag{19}$$

Refining the Model

It was evident from the data that, for structured interfaces, expert users (i.e. at t = 36) were, on average, able to achieve the complete task in one glance, even for the larger target arrays. This suggests that *pointing* seldom necessitated a separate, dedicated glance. Instead, it is suspected that participants began pointing (i.e. their hand left the steering wheel) during a *search/decide* glance, despite instructions to the contrary. To confirm this, we re-examined all individual glances made by participants for both structured and unstructured interfaces, specifically comparing the duration of the final glance with the predicted pointing component: in every situation, the pointing glance component (as defined by Equation 7), constituted only a proportion of the final glance (Figure 9). Thus, we conclude that the final glance constituted elements of both *search/decide* and *pointing*. This is perhaps unsurprising given that in a dual task/divided attention context, one would expect users to be naturally inclined to select a target as soon as it is located, rather than returning their attention to the primary task, as this would require them to relocate the target during a subsequent glance.

(Figure 9 about here)

We therefore modify our analysis approach, disregarding an isolated *pointing* element for NG, and assume that this glance is already included during the derivation of NG_{ds}. MGD can still be derived by dividing TGT (i.e. TGT_{pt} plus TGT_{ds}) by NG_{ds}, in line with International Standards definitions (ISO, 2014).

3.4. Combining Terms

Combining *searching*, *decision* and *pointing* terms, the following equations are therefore proposed:

$$TGT_{st} = \left(\frac{\log_2 N}{\log_2(N+t)}\right).(0.029N + 0.44) + 0.11\log_2 N + 0.11\log_2\frac{D}{W} + 0.35 \quad (20)$$

$$TGT_{un} = \left(\frac{\log_2 N}{\log_2(N+t)}\right).(0.10N - 0.028) + 0.049N + 0.045\log_2 N + 0.11\log_2\frac{D}{W} + 0.17 \quad (21)$$

$$NG_{st} = \left(\frac{\log_2 N}{\log_2(N+t)}\right).(0.021N + 1.04) + 1 \quad (22)$$

$$NG_{un} = \left(\frac{\log_2 N}{\log_2(N+t)}\right).(0.044N + 0.81) + 0.0071N + 1.96 \quad (23)$$

$$MGD_{st} = TGT_{st} \div NG_{st} \quad (24)$$

$$MGD_{un} = TGT_{un} \div NG_{un} \quad (25)$$

where:

st	=	structured
un	=	unstructured
N	=	total number of selectable items on the screen
t	=	number of exposures to interface
D	=	distance to target from hand position on steering wheel
W	=	target width

and $N > 1$

4. VALIDATION OF MODEL

Twelve people took part in the evaluation study: 5 male, 7 female. Mean age was 37.9, with ages ranging from 22 to 56 years. All participants held a valid driving licence and were experienced and active drivers (mean time with licence, 18 years; mean current annual mileage was 8500).

The device under evaluation was a novel touchscreen interface providing 'infotainment' and HVAC control functions. It was fully operational (with respect to the tasks under evaluation) and fitted to the driving simulator in the intended real world location within the centre console of the vehicle. Testing was conducted in accordance with the NHTSA Eye Glance Testing using a Driving Simulator (EGDS) test protocol (NHTSA, 2013) with an enhanced motorway scenario (as described in Section 2.2).

Three tasks were evaluated during the study; each began from a 'home' screen and required multiple screen interactions (Figure 10). Given that the tasks under investigation comprised multiple screen interactions, it was assumed that each separate screen/interaction constituted an isolated subtask (involving a 'search, find and select' activity) that could be modelled independently. For example, the 'HVAC' task (*change the airflow from 'balanced' to 'soft'*) constituted four 'subtasks': pressing the 'setting' button; pressing 'driver airflow'; pressing 'soft and quiet'; and pressing the 'close' button.

(Figure 10 about here)

In line with the NHTSA EGDS testing protocol, participants were provided with training on how to perform each testable task while stationary, how to drive the simulator while not performing a testable task, and how to perform each testable task while driving the simulated vehicle, before data collection began. As such, it could be assumed that participants were expert performers.

Following data collection, the proposed model (Equations 20 to 25) was used to predict visual demand for each task (defined by: TGT, NG and MGD metrics). During the calculations, the exposure measure (t) was selected as equal to 36, to reflect 'expert use' in accordance with our conjecture during the development of the model. It was also assumed that each subtask (and its associated metrics) occurred in series and thus predictions for each subtask could be aggregated to determine the total (TGT, NG) or mean (MGD) visual demand associated with each complete task. The theoretical predictions of visual demand were then compared to the empirically derived measures. To ensure thorough evaluation, results obtained for TGT, NG and MGD were initially compared to predictions made by both structured and unstructured models.

Figures 11, 12 and 13 show observed behaviour plotted against the structured and unstructured predictions for each task/subtask for TGT, NG and MGD, respectively. It is evident that, for all three tasks, the model overestimates TGT and NG for subtask 1 (i.e. interactions associated with the home screen). For subsequent subtasks, the observed behaviour generally falls between (or close to) the limits of structured and unstructured predictions (highlighted by shading in the Figures).

(Figure 11 about here)

(Figure 12 about here)

(Figure 13 about here)

5. GENERAL DISCUSSION

We describe the development of a model of human behaviour that predicts the visual demand associated with HMIs intended for use while driving. The approach draws upon the influential work by Cockburn et al. (2007), incorporating elements based on Hick-Hyman decision/search time and Fitts' Law pointing time. However, the model extends Cockburn et al.'s (2007) work by considering anticipation afforded by structuring and repeated exposure to an interface, and reflects the additional learning required to achieve expert status while using menus containing larger numbers of items; moreover, we aim to predict visual demand (rather than task time) in a dual task/divided attention situation, i.e. driving.

There are clear similarities between our model and the model proposed by Cockburn et al. (2007). A notable difference is the inclusion of an additional visual search/experience scale factor (d_{vs}), which reflects the increased learning required for larger target arrays, but also considers the additional visual demand due to primary/secondary task division/allocation. Indeed, drivers are required to divide their attention between driving (primary task) and interacting with the interface (secondary task) and may therefore be required to resume their search on multiple occasions, rather than only during initial exposure to the interface. Consequently, an element of 'searching' is likely to persist, even after multiple interactions with a well-designed, 'learnable' interface. In contrast, no such 'chunking' and search resumption would be expected in situations involving static menus in a sedentary context and thus the 'searching' element quickly decays through repeated exposure, evident within Cockburn et al.'s (2007) model.

A fundamental question posed by the research is whether the visual behaviour associated with secondary task execution while driving is amenable to theoretical analysis and modelling. Fitts' and Hick-Hyman theories are predicated on the ability to identify predictable patterns of behaviour. The current scarcity of Fitts and Hick-Hyman applications within driving-related HCI literature may therefore be indicative of concerns regarding the consistency and predictability of drivers' visual tendencies during secondary task execution. Indeed, there is natural variability regarding drivers' willingness and propensity to take their eyes off the road to engage with secondary tasks.

As such, drivers have been classified according to their natural gaze behaviour, with so-called 'long glancers' more inclined to take their eyes of the road for periods greater than 2.0 seconds (Donmez et al., 2010). Nevertheless, our research is predicated on the fact that interacting with an in-vehicle HMI presents an intrinsic and underlying visual demand that is fundamentally determined by the design of the interface and context of use. The model therefore assumes that a driver has decided (or is required) to interact with an in-vehicle HMI, and aims to predict the visual load that this demands. We are not attempting to model visual distraction, or how other driving-related factors (e.g. variations in primary task load, individual differences etc.) may influence visual engagement. The model therefore effectively aims to predict the minimal visual demand afforded by an HMI from a design perspective, and assumes that this is unaffected by the individual glance patterns or behaviour of test participants or drivers. As such, there is clear evidence from the data that, under these assumptions, single target acquisition using in-vehicle touchscreen HMIs constrains the user to predictable patterns of visual behaviour, suggesting that it is indeed amenable to theoretical analysis.

A possible limitation of the model derivation, however, is that it initially assumes that 'deciding/searching' and 'pointing' occur independently of one another (i.e. drivers complete the searching/locating task before commencing target selection), and thus the visual demand can be succinctly divided between these two activities. While such behaviour may be enforced or constrained experimentally (participants were instructed to maintain their hand on the steering wheel until they began pointing), aspects of searching and pointing may occur in parallel in real world applications, i.e. drivers may begin to 'point' before fully completing their 'search' activities. However, isolating these activities was necessary during the initial analysis, given Fitts and Hick-Hyman assumptions, and is in line with other, similar work, such as Cockburn et al. (2007). Further work is required to validate this approach.

An important consideration when designing HMIs for minimum visual demand (i.e. intended for in-vehicle placement) is the extent to which users are able to anticipate the location of their chosen option or function. Anticipation can be encouraged through repeated exposure and/or structuring items – repeated exposure enhances familiarisation and allows users to anticipate the location of target items, whereas structure can provide clues about target location. Structuring can be achieved by arranging options in alphabetical, numerical or chronological order. However, effective structuring within graphical user interfaces can be difficult to achieve, particularly if using visual iconography, typical of current in-vehicle touchscreen applications. Thus, one could consider any new interface to be 'unstructured' when encountered for the first time. When the interface becomes more familiar (for example, through repeated exposure), and users are able to anticipate the locations of target items, it may be more appropriate to consider the HMI as a 'structured' interface. By deriving equations for both structured and unstructured interfaces, we can therefore predict the range of performance that may be achieved from initial exposure, typical of novice users, to expert performance achieved through repeated use. Indeed, by comparing the observed and predicted behaviours during the validation study, it was evident that (with the exception of the home-screen, i.e. subtask 1) the empirical data generally existed between structured and unstructured predictions, with notable variability for different subtasks/measures (Figures 11, 12 and 13). It could be concluded that in situations where the observed behaviour was more closely aligned with 'unstructured' predictions, users were less familiar with that particular subtask or screen layout; moreover, one may predict that visual behaviour would migrate towards 'structured' predictions as familiarity/expertise increases. From this perspective, the validation results suggest that users may have been more familiar with some screens than others – the observed visual demand associated with the home-screen in particular was significantly lower than even the structured predictions.

An alternative explanation is that the observed behaviour may be indicative of the effectiveness of the HMI design. Indeed, participants who took part in the validation study were trained in accordance with NHTSA EGDS protocol (NHTSA, 2013) and consequently were deemed to be expert users. Thus, in situations where the observed behaviour was more closely aligned with structured predictions, one could conclude that the interface or interactions associated with that task or subtask were designed well and afforded anticipation/efficient use. In contrast, visual behaviour tending towards the unstructured predictions may suggest that the interface or task lacked learnability (e.g., the interface was cluttered, target items were poorly located etc.). By deconstructing each task into component subtasks (each representing a single target interaction with the HMI), it is therefore possible to draw conclusions regarding the design of the interface (associated with each subtask) by comparing the observed performance of each subtask with its predicted behaviour (Figures 11, 12 and 13). Thus, one could conclude that in situations where the observed performance matches the 'structured' predictions, the HMI design encouraged efficient interactions and could be considered as 'well-designed'. In contrast, visual behaviour that is more closely aligned with the 'unstructured' predictions, may indicate poor design. Consequently, the predictive model also has utility as a formative design evaluation tool. However, it is noteworthy that the evaluation results were predicated on the notion that the tasks could be broken down into discreet subtasks, each of which could be modelled independently; this is not necessarily the case in all situations. Further work is required to explore this assumption.

An additional consideration is that assigning model parameters (e.g. number of items in an interface, target width, distance to target etc.) to real-world interfaces may be complicated by novel design techniques, such as skeuomorphic elements that may not lend themselves to theoretical analysis (i.e. complications in defining target width etc.). During the validation work, all interactions involved finger touch input using discrete and delineated elements; this may not be the case for all real-world interfaces. Therefore, further work that considers different design elements and interaction techniques is required.

It is also noteworthy that despite our model's apparent utility as both a design and evaluation tool, the resulting predictions are likely to be highly contextual. Testing was conducted in a medium fidelity, fixed-based driving simulator using a generic motorway scenario. If a different simulated driving scenario had been used during testing, some aspects of visual behaviour may have been affected (Large et al, 2015), and thus the derived equations would have differed. For the decision/search component, visual demand may also have been influenced by semantic complexity associated with the search tasks (although efforts were made to mitigate this affect by selecting unambiguous targets). Furthermore, in a real-world environment, other factors inherent with the in-vehicle environment (e.g. limited space, restricted movement/anthropometry, location of touchscreen, arm instability, roadway vibrations, enforced operation using non-dominant hand in right-hand drive vehicle etc.), may also influence secondary task visual demand while driving (particularly NG and MGD) and the simulated vehicle, experimental approach and participant cohort may have been insufficient to fully represent all factors.

It is also noted that the derived models are complex, requiring up to five parameters to be fitted to data to predict performance. Given this complexity, some factors, such as the numbers of participants used to derive the model, may be considered relatively low. Consequently, absolute predictions of visual demand should be treated with caution and derived data should serve as a guide only (further validation work is ongoing). Nevertheless, the model still has genuine utility in its ability to predict relative metrics, e.g. comparing several prospective design concepts early in the design cycle, thereby reducing arguments to simple calculations based on an understanding of the underlying characteristics of the HMI and task (Raskin, 2000).

6. CONCLUSIONS AND FUTURE WORK

We describe the development of a predictive model of visual behaviour associated with in-vehicle HMIs. Empirical testing and initial validation suggest that the derived model may be capable of making valuable predictions regarding the visual demand presented by such interfaces. Furthermore, it possesses utility as both a design and evaluation tool. In line with similar theoretical work (e.g. Cockburn, 2007), the model assumes that interactions with an HMI can be considered as comprising separate decision/search and pointing components that are amenable to Fitts' Law and Hick-Hyman theory: this assumption, as well as other limitations, such as the relatively low number of participants who took part in each study and the model's un-tested application to multiple-screen interactions, need to be revisited in future investigations.

Although the research focused on single finger-touch pointing tasks using a touchscreen while driving, the results are not limited to driving and can be generalized to any setting where users interact with an interface, making a choice decision followed by a pointing task (i.e. 'search, find and select' tasks), while concurrently executing a more critical primary task. However, empirical work would be required to derive and validate models specific to other situations and contexts. Further work should also consider the visual demand of more complex interactions, e.g. surface gestures, as well as investigating other techniques, such as grouping UI elements and skeuomorphism, intended to enhance HMI usability, learning and appeal.

7. ACKNOWLEDGEMENTS

The research was conducted in collaboration with Jaguar Land Rover Research, with validation work supported by Jaguar Land Rover. The authors would like to gratefully acknowledge their support.

Figure 1. Driving simulator used during studies showing motorway scenario

Figure 2. Experimental set-up showing (clockwise from top left) (i) touchscreen with target, (ii) participant wearing ETG, (iii) screen located in 'lower' position, (iv) ETG visual trace

Figure 3. Fitts' 'pointing' relationship for TGT during single target acquisition while driving

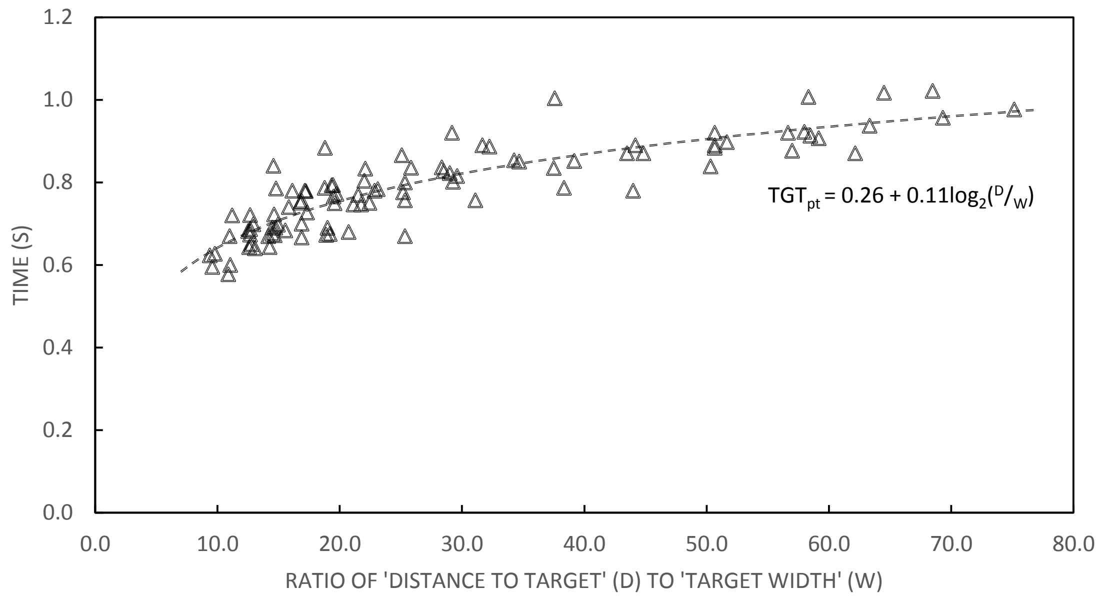


Figure 4. Target layouts used during study two

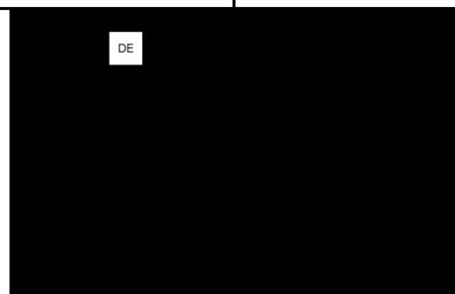
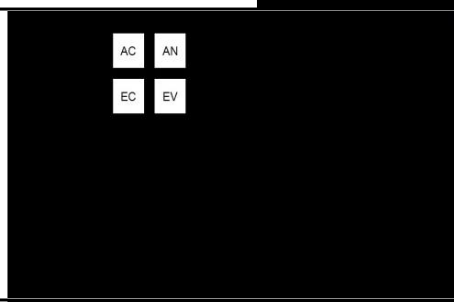
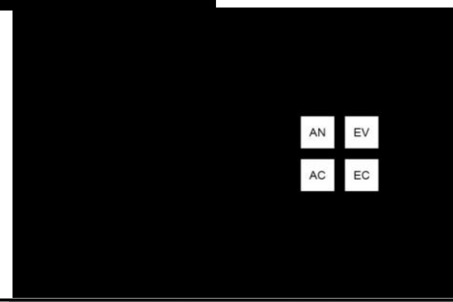
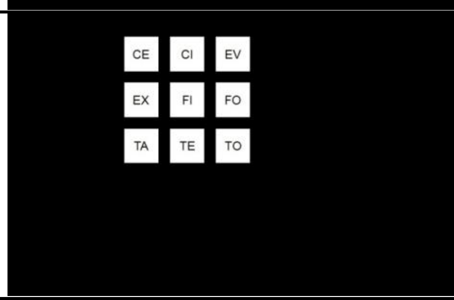
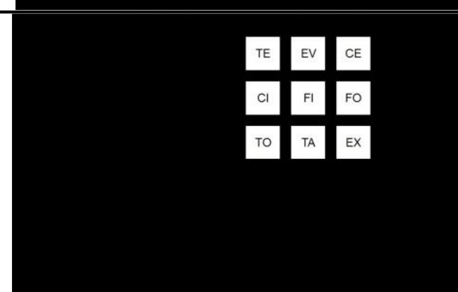
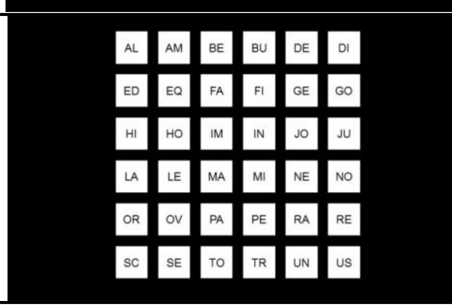
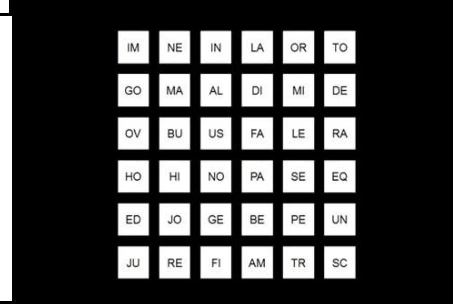
	Alphabetically Structured (row-by-row)	**Unstructured**
N=1 1x1 array Target word: *Development*		
N=4 2x2 array Target word: *Economic*		
N=9 3x3 array Target word: *Example*		
N=36 6x6 array Target word: *November*		

Figure 5. Fitts' pointing relationship for TGT showing observed behaviour for N={1, 4, 9, 36} and derivation of additional pointing glance time modifier

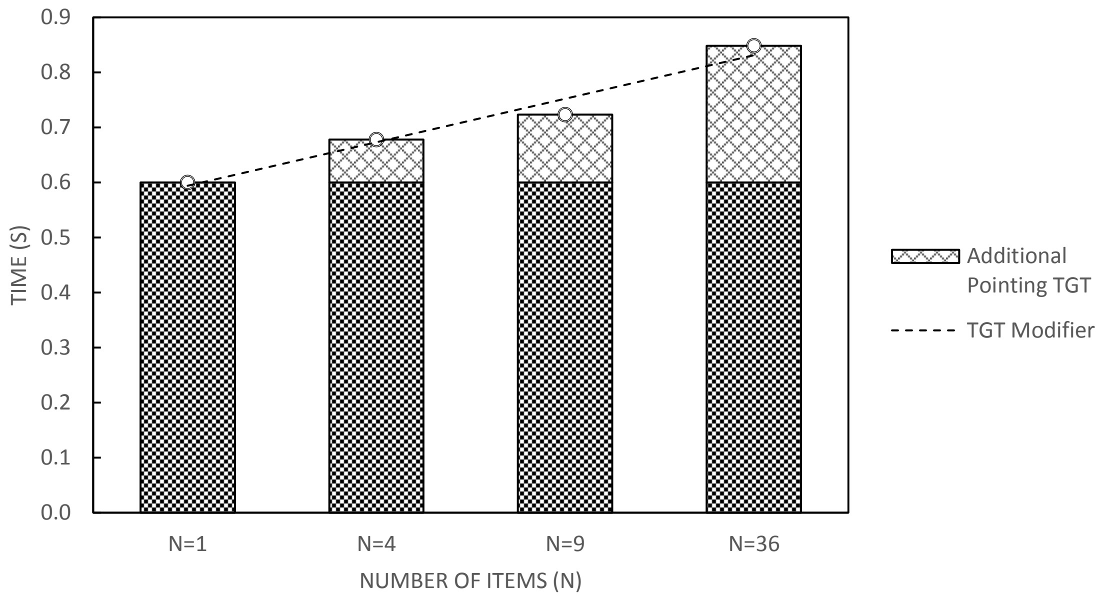


Figure 6. Modelling visual search behaviour: TGT @ t=1 for structured and unstructured arrays

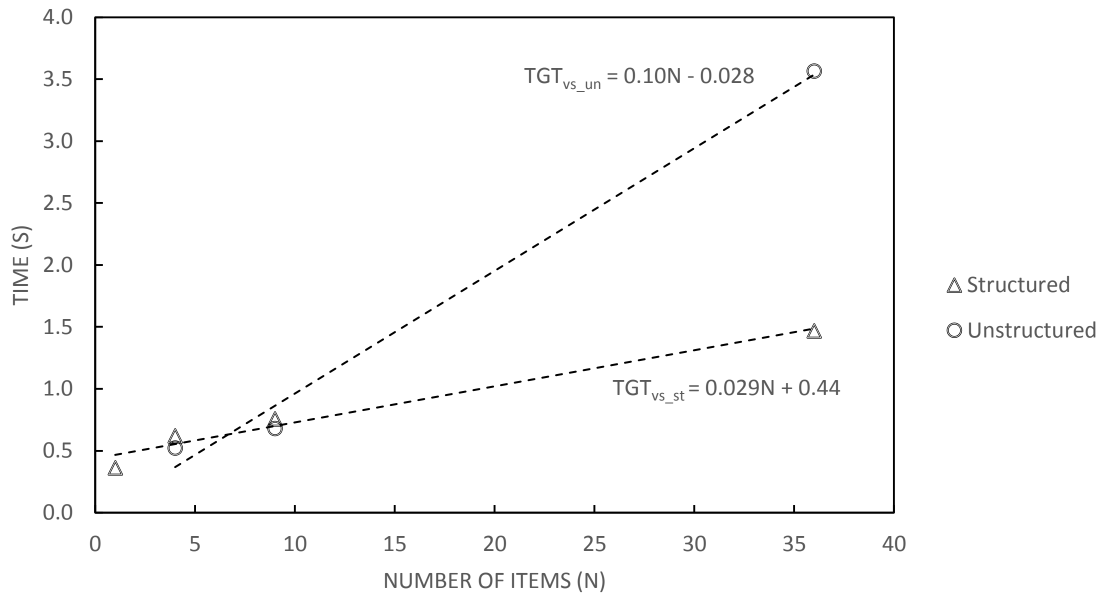


Figure 7. Modelling visual search behaviour: NG @ t=1 for structured and unstructured arrays

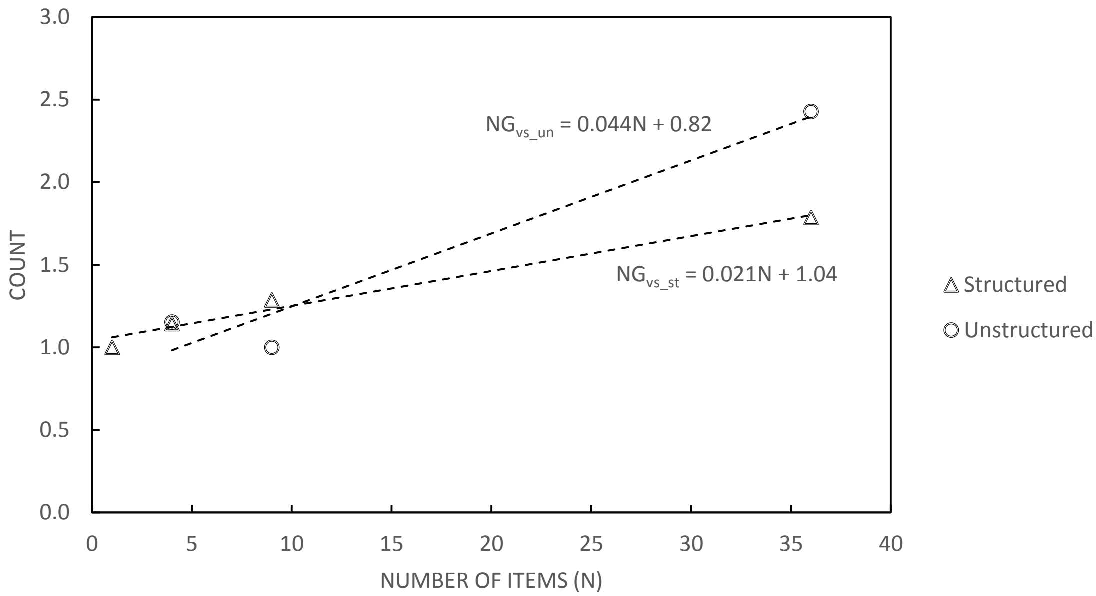


Figure 8. Modelling Hick-Hyman decision behaviour: TGT @ t=36 for structured and unstructured arrays

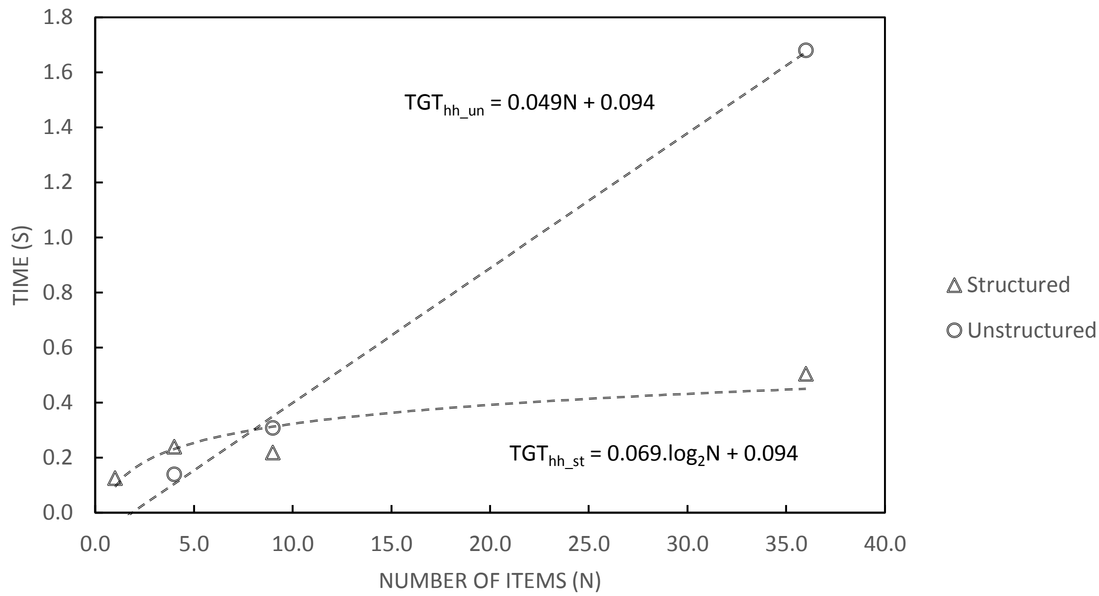


Figure 9. Visualisation of final glance, highlighting 'decision/search' (darker shading) and 'pointing' (lighter shading) components

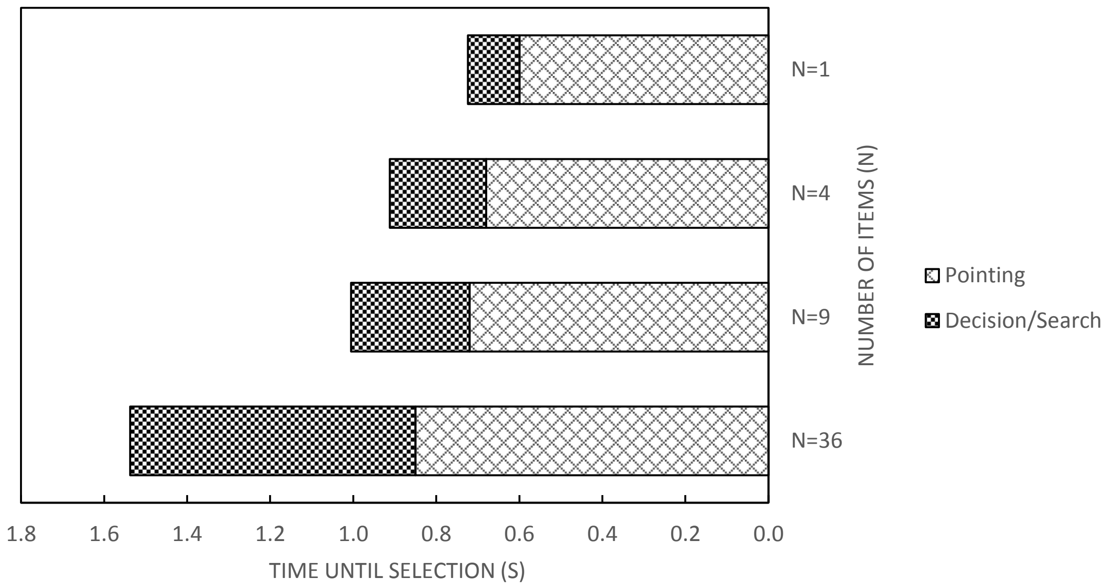


Figure 10. Details of evaluation tasks

Task	Description	No. of subtasks
Entertainment	Change the listening mode to ‘radio’	3
HVAC	Change the airflow from ‘balanced’ to ‘soft and quiet’	4
Personal Comfort	Change the driver seat massage mode to ‘shoulder’ and set the massage intensity to ‘level 4’	4

Figure 11. Structured and unstructured model predictions of TGT compared to observed performance (isolated by subtask) showing Entertainment (E), HVAC (H) and Personal Comfort (C) tasks

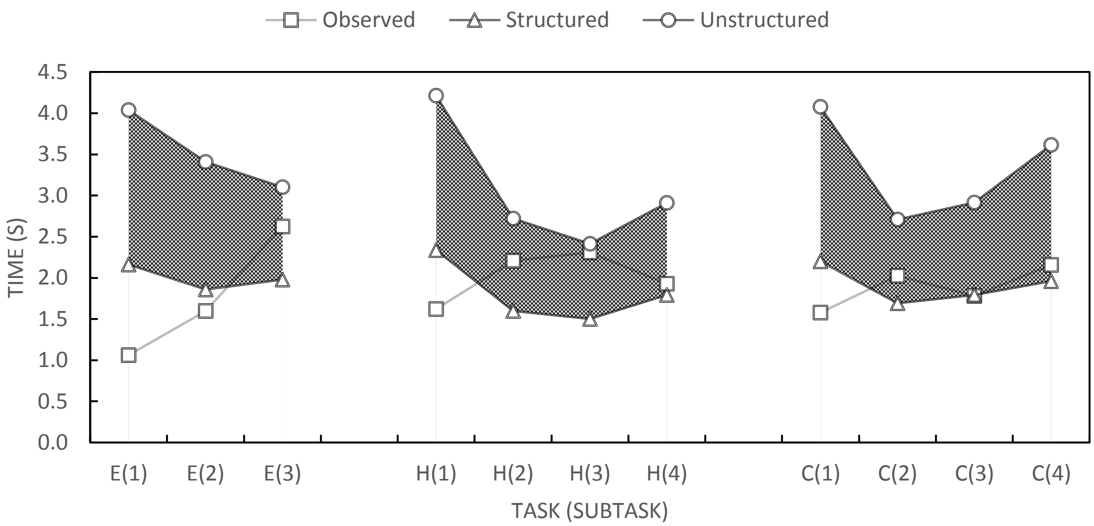


Figure 12. Structured and unstructured model predictions of NG compared to observed performance (isolated by subtask) showing Entertainment (E), HVAC (H) and Personal Comfort (C) tasks

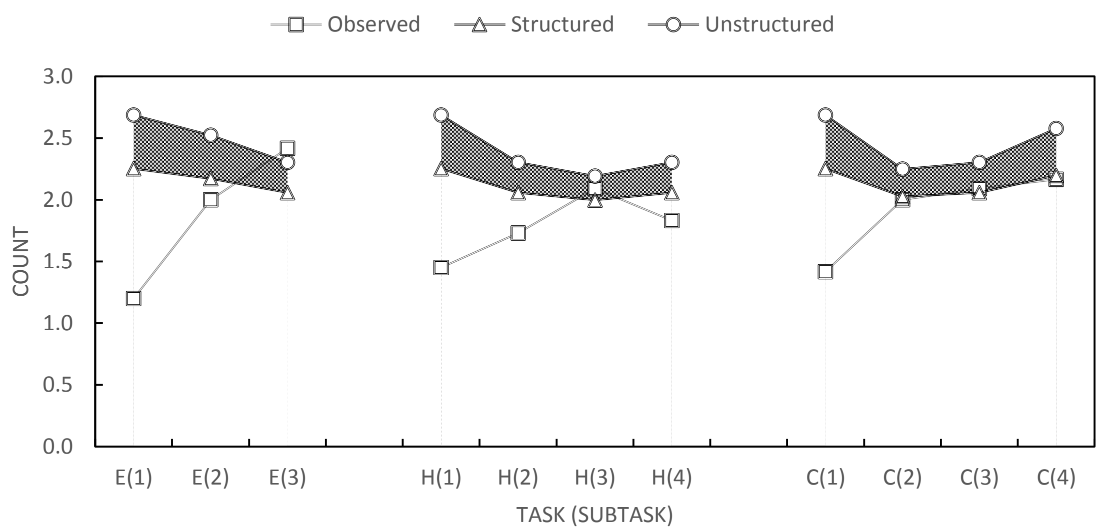
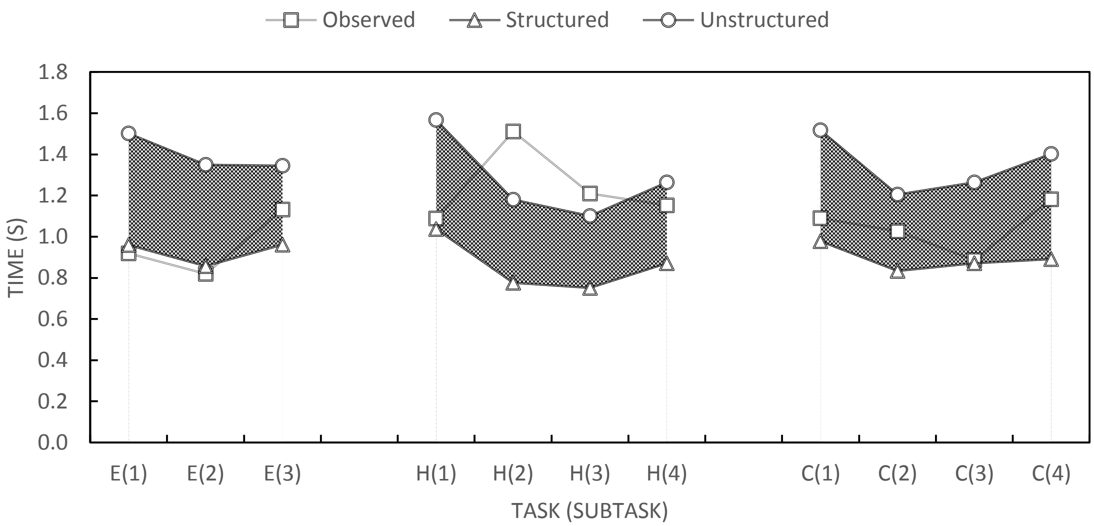


Figure 13. Structured and unstructured model predictions of MGD compared to observed performance (isolated by subtask) showing Entertainment (E), HVAC (H) and Personal Comfort (C) tasks



REFERENCES

Bi, X., Smith, A. B., Zhai, S. 2012. Multilingual Touchscreen Keyboard Design and Optimization. *Human-Computer Interaction*, 27 (4), 352 - 382.

Beggs, W.D.A., Graham, J.C., Monk, T.H., Shaw, M.R.W. and Howarth, C.I., 1972. Can Hick's law and Fitts' law be combined?. *Acta psychologica*,*36*(5), pp.348-357.

Burnett, G., Lawson, G., Millen, L., Pickering, C. 2011. Designing touchpad user-interfaces for vehicles: which tasks are most suitable? *Behaviour & Information Technology*, 30(3), 403-414.

Card, S. K., English, W. K., Burr, B. J. 1978. Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics*, 21(8), 601-613.

Cockburn, A., Ahlström, D., Gutwin, C. 2012. Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse. *International Journal of Human-Computer Studies*, 70(3), 218-233.

Cockburn, A, Gutwin, C., Greenburg, S. 2007. A predictive model of menu performance, *Proceedings of SIGCHI conference*, ACM: NY.

Donmez, B., Boyle, L.N., Lee, J.D., 2009. Differences in off-road glances: Effects on young drivers' performance. *Journal of transportation engineering*, 136(5), pp.403-409.

Fitts, P. M. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 47(6), 381.

Harvey, C., Stanton, N.A., Pickering, C.A., McDonald, M. and Zheng, P., 2011. In-vehicle information systems to meet the needs of drivers. *Intl. Journal of Human–Computer Interaction*, *27*(6), pp.505-522.

Hick, W.E., 1952. On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*(1), pp.11-26.

Hoffmann, E.R., Lim, J.T., 1997. Concurrent manual-decision tasks. *Ergonomics*, *40*(3), pp.293-318.

Hyman, R., 1953. Stimulus information as a determinant of reaction time. *Journal of experimental psychology*, *45*(3), p.188.

ISO 15007-1:2014. 2014 *Road vehicles – Measurement of driver visual behaviour with respect to transport information and control systems – Part 1: Definitions and parameters*

John, B.E., Salvucci, D.D., Centgraf, P. and Prevas, K.C., 2004, July. Integrating Models and Tools in the Context of Driving and In-vehicle Devices. In *ICCM* (pp. 130-135).

Kim, I. and Jo, J.H., 2015. Performance Comparisons Between Thumb-Based and Finger-Based Input on a Small Touch-Screen Under Realistic Variability. *International Journal of Human-Computer Interaction*, *31*(11), pp.746-760

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., Ramsey, D. J. 2006. *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data* (No. HS-810 594).

Landauer, T.K., Nachbar, D.W., 1985. Selection from alphabetic and numeric menu trees using a touch screen: breadth, depth, and width. *ACM SIGCHI Bulletin*, *16*(4), pp.73-78.

Large, D.R., van Loon, E., Burnett, G. and Pournami, S., 2015. Applying NHTSA task acceptance criteria to different simulated driving scenarios. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 117-124). ACM.

Lewis, J. R., Kennedy, P. J., & LaLomia, M. J. 1999. Development of a Diagram-Based Typing Key Layout for Single-Finger/Stylus Input. *Proc. of The Human Factors and Ergonomics Society 43rd Annual Meeting*.

MacKenzie, I. S. 1992. Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction*, 7(1), 91-139.

NHTSA. 2013. *Visual-Manual NHTSA Driver Distraction Guidelines For In-Vehicle Electronic Devices.* NHTSA-2010-0053

Pettitt, M., Burnett, G. and Stevens, A. 2007. An extended keystroke level model (KLM) for predicting the visual demand of in-vehicle information systems, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*

Raskin, J., 2000. *The humane interface: new directions for designing interactive systems*. Addison-Wesley Professional.

Salvucci, D.D., 2001. Predicting the effects of in-car interface use on driver performance: An integrated model approach. *International Journal of Human-Computer Studies*, *55*(1), pp.85-107.

Seow, S.C., 2005. Information theoretic models of HCI: a comparison of the Hick-Hyman law and Fitts' law. *Human-Computer Interaction*, *20*(3), pp.315-352.

Shannon, C. E., Weaver, W. 1949. *The mathematical theory of information*.

Soukoreff, R.W., MacKenzie, I.S., 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *International journal of human-computer studies*, *61*(6), pp.751-789.

AUTHOR BIOGRAPHIES

David R Large

Dr David R. Large is a Research Fellow with the Human Factors Research Group at the University of Nottingham. He has been involved in a broad range of projects concerning the design/acceptance of novel in-vehicle interfaces and systems, and the development and application of experimental techniques to support this work.

Gary Burnett

Dr Gary Burnett is an Associate Professor in Human Factors at the University of Nottingham. He has over 20 years' experience in Human Factors research. His work addresses key safety, usability and acceptability issues for advanced in-car systems, and he is particularly concerned with the assessment of driver distraction.

Elizabeth Crundall

Dr Lizzie Crundall is a Human Factors research consultant with over 15 years' experience in experimental design and statistical analysis, specialising in eye-tracking and simulation methodologies. Her most recent projects include studies of motorcyclist behaviour, hazard perception, road sign design, and the effects of HMI attributes on driver distraction.

Editha van Loon

Dr Editha van Loon is a Research Fellow in the School of Psychology at the University of Nottingham. With over 15 years' experience in research, she holds particular expertise in understanding the effect of factors such as age, driving experience and pre-existing clinical conditions on the visual behaviour of drivers.

Ayse Leyla Eren

Ayse is a PhD Postgraduate Researcher at the University of Nottingham. Her research is concerned with reducing the visual demand of in-vehicle touchscreen HMIs, with a focus on developing guidelines to support novel interface designs/interaction techniques that promote non-visual interaction, for example by encouraging peripheral vision as an interaction mechanism.

Lee Skrypchuk

Lee works at Jaguar Land Rover in the Research Department. As a Human Machine Interface Technical Specialist, Lee focuses on technical aspects such as Engineering, Human Factors and Psychological. Lee has led projects including Head-Up Display, Driver Monitoring and Gesture control, and holds responsibility for research roadmaps in this area.