A Predictive Model of the Visual Demand Associated with In-Vehicle Touchscreens

D.R. Large[1*], G. Burnett[1], E. Crundall[1], E. van Loon[1], A. L. Eren[1], L. Skrypchuk[2]

[1]Human Factors Research Group, Faculty of Engineering, University of Nottingham, UK
[2]Jaguar Land Rover Research, International Digital Laboratory, Coventry, UK
[*]david.r.large@nottingham.ac.uk

Abstract: Touchscreen-HMIs are increasingly popular within vehicles. Understanding the likely visual demand of new designs is therefore important but typically requires time-consuming and costly testing with functioning prototypes. Theoretical modelling allows performance to be determined much earlier in the design cycle, but has seldom been applied to touch-screen interfaces in divided-attention contexts, such as driving. We describe a theoretical model of human performance – derived from empirical testing – that makes *a priori* predictions of the visual demand (total glance time, number of glances and mean glance duration) elicited by finger-touch pointing tasks in a driving context. The model integrates two well-established laws of human behaviour – the Hick-Hyman Law, concerning decision/search behaviour, and Fitts' Law, which considers the movement to acquire a visual target. The model also recognises that menus with greater depth will extend decision/search time and delay the time taken to achieve expert status. Preliminary validation work, comparing predictions for a real-world prototype touchscreen interface with empirically-obtained data, suggests that the model may provide an effective design and evaluation tool capable of making valuable predictions regarding the limits of visual demand/performance associated with in-vehicle interfaces, enabling designers to explore a wide range of possible designs before implementation, and permitting cost-effective redesign. Further work is required to refine the model, particularly in consideration of more complex tasks, involving multiple screen interactions.

1. Introduction

Touchscreens Human-Machine Interfaces (HMIs) are increasingly becoming the primary display and control interface in vehicles. However, they inherently demand some visual attention, often relying on visual cues, in lieu of tactile prompts, and this can divert a driver's attention away from the road scene. Determining the amount of visual attention demanded by such interfaces is therefore important but this traditionally necessitates empirical testing and user trials (e.g. [1]) requiring fully-functioning prototypes. Such testing can also be time-consuming and costly to conduct, and may result in expensive re-design and re-evaluation. Theoretical modelling allows human performance to be determined much earlier in the design cycle than traditional design-evaluation techniques and therefore reduces arguments to simple calculations, based on an understanding of the underlying characteristics of the interface and task [2]. This enables designers to explore a wide range of possible designs before implementation, and permits cost-effective redesign. Although there has been significant interest in the theoretical prediction of human behaviour and performance within the field of HCI, this has typically focussed on technology and interfaces that act as the only or primary focus for a user's attention (e.g. menu navigation using desktop computers), with the aim of predicting interaction time or performance, and has seldom been applied to interfaces in divided-attention contexts, such as driving (although there have been some notable forays, such as Pettitt et al. [3], who extended key-stroke level modelling (KLM) to account for the context of driving).

1.1. Information Theory

A common basis to model human performance is Information Theory [4], which views humans as information processors. Adaptations of Information Theory were first applied to HCI during the 1980s by Card et al. [5], who articulated two information theoretic models as guiding principles to enhance technology and interface design and usability, notably Fitts' Law [6] and the Hick-Hyman Law [7, 8].

Fitts' Law is concerned with predicting the time taken to move to an item 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 activity. It predicts a logarithmic relationship between movement time and target-distance and width because human pointing to visual targets typically allows a large proportion of the distance towards the target to be completed rapidly without attending to feedback. Essentially, Fitts' Law predicts that small, distant targets require more information processing than larger, closer targets, and therefore take longer to acquire.

The Hick–Hyman Law 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. The law predicts a linear relationship between reaction time and the item's information content (e.g. number of options on a menu); it also predicts that as users become more experienced, their visual search strategy/decision time will change from linear to logarithmic, i.e. the time taken to locate an item reduces significantly as familiarity increases.

Both Fitts' Law and the Hick-Hyman Law are highly applicable within the fields of HCI and Human Factors. However, much of the proceeding work has focussed largely on the application of Fitts' Law in isolation, therefore overlooking the time taken to choose and locate the target, or has failed to consider adaptations to visual search strategies and any associated changes in search time as users become more familiar with interfaces. Moreover, each law has tended to be applied in isolation.

Combining Fitts' and Hick-Hyman laws can provide a more expansive prediction of human behaviour than each model offers alone (for examples see: [9, 10, 11]), but these have typically achieved limited success. More recently, Cockburn et al. [12] proposed a model of menu performance – that combines element of Fitts' and Hick-Hyman laws – to predict the time to find a target item and the time to select (or move to) that item; the model also recognises users' increasing expertise, reflected in a gradual move from a linear to logarithmic visual search/decision time. Results of the validation studies conducted by Cockburn et al. [12] suggest that their predictions of static task time were generally very accurate – within 2% of empirically collected data. However, a significant limitation of the Cockburn et al. [12] model – at least from our own research perspective – is that it only applies to the prediction of static task time in a sedentary context. As automotive ergonomists, we are interested in whether similar predictions could be made regarding interactions with in-vehicle HMIs while driving: to date, there is no evidence of successful attempts to combine Fitts' and Hick-Hyman Laws to model 'search, find and select' interactions with a touchscreen HMI in a driving context.

1.2. Approach and Assumptions

In a driving context, vision provides the primary source of information available to drivers. Therefore, while total task-time is commonly used an indicator of the suitability of secondary tasks or devices, the visual attention demanded by such tasks is likely to be a far better predictor of suitability [1]. Moreover, 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 [13, 1].

Secondary task visual demand is typically measured using three key metrics – total glance time (TGT), mean glance duration (MGD) and number of glances (NG) – with guidelines recommending 'safe' limits associated with each metric. In accordance with ISO 15007 part 1 [14], TGT is defined 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 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" [14].

In line with Cockburn et al. [12], we consider 'search, find and select' tasks to comprise *decision/search* (Hick-Hyman) and *pointing* (Fitts) components. Cockburn et al. [12] were concerned with predicting the time to select an item using static menus. Assuming similar logic, we predict that the glances associated with selecting an item, *i*, comprises glances associated with both *deciding/searching (ds)* and *pointing (pt)*. Thus,

$$TGT_i = TGT_{ds} + TGT_{pt} \quad (1)$$

$$NG_i = NG_{ds} + NG_{pt} \quad (2)$$

For MGD, we assume that this can be obtained from TGT and NG data, in line with International Standards definitions [14].

$$MGD_i = TGT_i \div NG_i \quad (3)$$

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; anything prior to this action (following the presentation of a new stimulus/instruction) was deemed to be associated with *deciding/searching*.

2. Method

Derivation of the model incorporated three studies, each conducted independently. Common elements, such as the driving simulator and set-up/approach, are summarised below, with further study-specific details provided in the subsequent sections. Study one considered the pointing component of single-target acquisition, thereby deriving an initial Fitts' Law relationship. Study two considered the effect of additional flanking items on pointing efficiency – thereby modifying the relationship derived in study one – and isolated an associated Hick-Hyman decision/search component. Study three provided an initial validation of the derived model by applying it to a real-world prototype touchscreen interface.

2.1. Apparatus, Design and Procedure

All testing took place using a medium-fidelity, fixed-based driving simulator based at the University of Nottingham. Participants comprised experienced and regular drivers who responded to advertisements posted online and around the University campus; they were reimbursed with shopping vouchers as compensation for their time and provided written informed consent before taking part.

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).

Fig. 1. *Driving simulator showing motorway scenario used during stuides*

The car-following dual task paradigm was employed throughout testing. This approach is typically employed in driver distraction research to control primary task workload, and is specified as part of a standardised experimental protocol within international driving standards (e.g. [1]).

During each study, stimuli were displayed on an HP EliteBook 2740p tablet computer that was located in a representative location within the centre console of the car. During study one, the location of the tablet was alternated between two different positions – notionally referred to as ‘upper’ and ‘lower’ – in order to capture a larger range of ‘distances to target’ (Fitts’ law ‘D’ metric). In all other testing, the tablet computer was located in the ‘upper’ position – a more common location for an in-vehicle, centre console display (Figure 2). 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; eye-tracking data were subsequently coded using semantic gaze mapping. Secondary task/response time, comprising search time and pointing time, was also recorded. Forty experienced and active drivers took part in testing (23 male, 17 female, with a mean age of 31.4 years; mean annual mileage was 6390), with a different cohort recruited for each of the separate studies.

Fig. 2. *Location of touchscreen within driving simulator buck (example taken from study 1)*

2.2. Study 1: Fitts' Pointing Component (Single Targets)

During study one, participants undertook two counterbalanced driving sessions – one with the touchscreen located in the upper position, the other, with it in the lower position. 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 'distance to target' (required for Fitts' law) could be determined for each location.

For each stimulus, a single square target item appeared on the touchscreen (Figure 2), accompanied by an audible tone to inform participants of the presence a new target. Participants were instructed to touch ('point at') the target as quickly and accurately 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, following a short delay, a new target appeared. Targets varied in size (6, 12, 18, 24mm) and location, based on existing in-vehicle HMI guidelines. The order of presentation of target locations and dimensions was randomised between participants.

2.3. Study 2: Fitts' Pointing Component (Flanking Targets) and Hick-Hyman Decision/Search Component

In order to add a decision/search element, participants were required to find and select a single target item located amongst an array of similar items during study two. 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. All words conformed to regular UK English phonetic pronunciation and spelling and were between 6 and 12 letters in length.

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; target size was consistent throughout the study, in line with existing in-vehicle HMI guidelines, and was the median size used during study one (i.e. 15mm). Targets were presented as either alphabetically-structured arrays, affording anticipation, or unstructured arrays (Figure 3). Participants undertook seven drives (i.e. there was only one drive for N=1), each of which constituted a different array size and structure, and lasted approximately 5 minutes. 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 a target appearing in each of the possible locations. After selecting the correct target, the array disappeared from the screen and, following a short delay, the array re-appeared, preceded by a new auditory cue. 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.

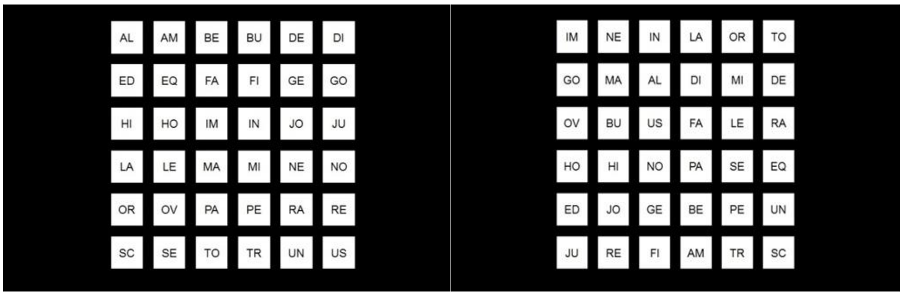


Fig. 3. *Example of 6x6 arrays showing structured (left) and unstructured (right)*

2.4. Study 3: Model Validation

As a preliminary validation, a novel touchscreen interface providing 'infotainment' and HVAC control functions was evaluated in accordance with the NHTSA Eye Glance Testing using a Driving Simulator (EGDS) test protocol [1]. Three tasks were evaluated: changing the listening mode to 'radio' ('Entertainment'), changing the airflow from 'balanced' to 'soft and quiet' ('HVAC') and changing the driver seat massage mode to 'shoulder' and increasing the massage intensity to 'level 4' ('Personal Comfort'). Following the completion of empirical testing, the derived model was used to predict the visual demand of each task (TGT, NG and MGD). The theoretical predictions of visual demand were then compared to the empirically derived measures.

3. Results and Analysis

The results of study one were used to determine the Fitts' pointing component for single targets. For the purpose of analysis, it is assumed that there is no 'search' associated with single target acquisition.

3.1. Study 1: Fitts' Pointing Component – Single Targets

Pointing behaviour is derived from Fitts' Law. Results show a strong linear relationship between Fitts' 'index of difficulty', $\log_2 \frac{D}{W}$ (where D = distance to target and W = width of target), and pointing TGT, suggesting that Fitts' law applies to single target acquisition in a driving context (Figure 4).

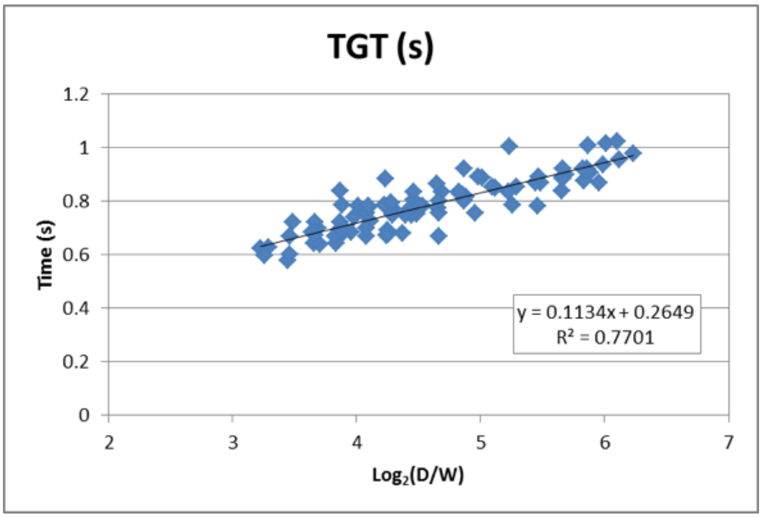


Fig. 4. *Fitts' Law relationships derived from study 1*

The following relationship is thus derived:

$$TGT_{pt} = 0.26 + 0.11 \log_2 \frac{D}{W} \; (R^2 = 0.77) \quad (4)$$

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

3.2. Study 2: Fitts' Pointing Component – Flanking Targets

During study two (where 'search' activities were also required during most tasks), we assumed that pointing began when the participant's hand left the steering wheel, and instructed participants as such during testing. In situations where single items were presented, predictions of TGT_{pt} corresponded well with predictions from study one. However, in the presence of additional flanking items, it was evident that TGT_{pt} exceeded the predictions from the first study, suggesting that pointing efficiency was degraded in the presence of additional targets, thereby increasing visual demand (Figure 5). The additional visual demand encouraged by the presence of multiple targets can be derived empirically as: $0.045\log_2 N$ $(R^2 = 0.98)$, where N = the number of targets. Thus, we appended this term to Equation 4.

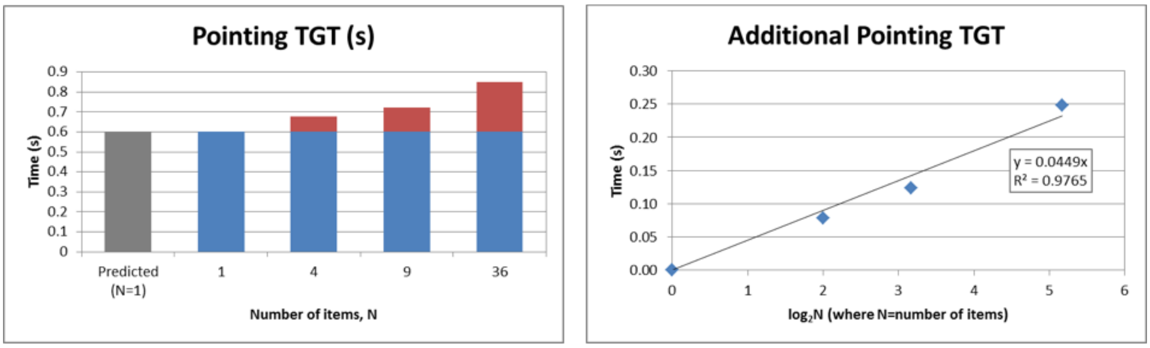


Fig. 5. *Fitts' pointing relationship for TGT, showing: observed versus predicted behaviour from study one (red bars=additional pointing time) (left), and derivation of additional pointing glance time modifier due to flanking targets (right)*

$$TGT_{pt} = \left(0.26 + 0.11\log_2\frac{D}{W}\right) + (0.045\log_2 N) \quad (5)$$

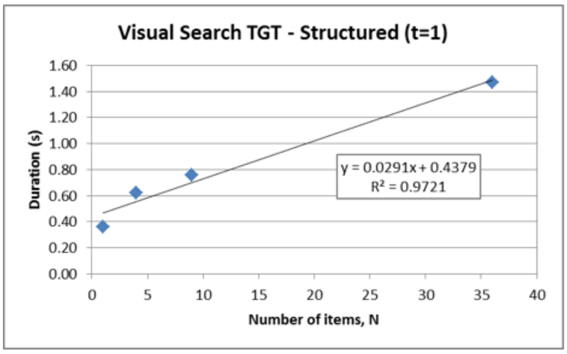
Again, we assumed that *pointing* requires only one glance, so NG_{pt} remained equal to one and MGD_{pt} could effectively be determined using the same expression as TGT_{pt}.

3.3. Study 2: Fitts' Hick-Hyman Search Component

In line with Cockburn et al. (2007), we assumed that decision/search behaviour (in the presence of additional, flanking targets) can be determined by interpolating behaviour between a linear visual search component (*vs*) and a logarithmic Hick-Hyman decision component (*hh*).

However, given that our model concerns the prediction of visual behaviour in a driving context, it is recognised that the visual attention directed towards a secondary interface/task will naturally be restricted by the demands of the primary driving task, and further likely to be interrupted as drivers are required to routinely attend to the primary driving task. Thus, rather than 'user expertise' *per se* (as defined by Cockburn et al. [12]), we expect decision/search behaviour to be influenced by a driver's ability to *anticipate* the location of the target and *resume* searching after their attention has been directed elsewhere (e.g. towards the driving task). Consequently, our evaluations of decision/search activities associated with multiple targets also concerned structured (*st*) and unstructured (*un*) arrays.

Moreover, we assumed that when a user is inexperienced, they are unable to anticipate the target location, and thus there is a linear relationship between TGT and the total number of items – for both structured and unstructured interfaces/arrays – in line with Cockburn et al.'s [12] approach. Plotting these data for inexperienced users (i.e. at t=1, where t = the number of exposures to an interface), we confirm a linear relationship for TGT associated with both structured and unstructured arrays (Figure 6). A similar linear relationship can be derived for NG (no figure).



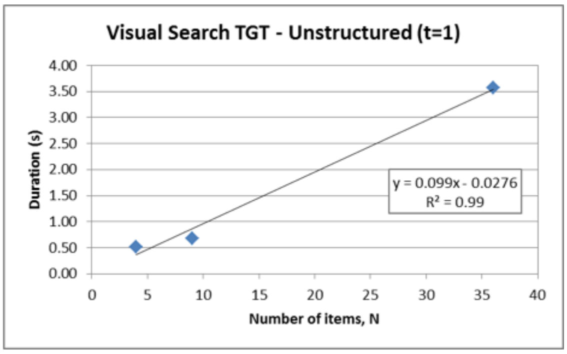


Fig. 6. *Modelling visual search behaviour – TGT @ t=1 for structured (left) and unstructured (right) interfaces, where t = the number of exposures*

The following relationships are thus derived:

$$TGT_{vs\_st} = 0.029N + 0.44\ (R^2 = 0.97) \tag{6}$$

$$TGT_{vs\_un} = 0.10N - 0.028\ (R^2 = 0.99) \tag{7}$$

$$NG_{vs\_st} = 0.021N + 1.04 \; (R^2 = 0.98) \tag{8}$$

$$NG_{vs\_un} = 0.044N + 0.81 \; (R^2 = 0.94) \tag{9}$$

Further to this, it is assumed that all HMIs intended for deployment within vehicles are ultimately 'learnable', as defined by Cockburn et al. [12] (i.e. items remain in fixed locations). It is also expected that users will require more trials to achieve 'expert' status when using menus containing multiple targets. However, given the divided-attention context of driving (i.e. drivers may be required to stop and resume search activities on several occasions – as dictated by the demands of the primary driving task – particularly for larger arrays), it is expected that the *search* component, e.g. TGT_{vs}, will persist longer in the presence of additional target items compared to the predictions made by Cockburn et al. [12]. Using the empirical data, a visual search/experience scale factor, d_{vs}, is thus derived (Equation 10), applicable to situations of divided attention. Consequently, Equation 1 can be redefined as Equation 11, below.

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

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

3.4. Study 2: Fitts' Hick-Hyman <u>Decision</u> Component

As users become more experienced with an interface, they are able to anticipate the location of their target item based on spatial memory. Using the data for experienced users (i.e. at t=36, taken as the maximum exposure recorded during the studies and the point at which users were therefore expected to be familiar with the array), a logarithmic Hick-Hyman relationship can be derived. For *structured* interfaces, the linear visual search term is thus replaced by a logarithmic Hick-Hyman relationship:

$$TGT_{hh\_st} = 0.069 \log_2 N + 0.094 \; (R^2 = 0.84) \tag{12}$$

In contrast, the data obtained from exposures to *unstructured* interfaces suggest that the relationship between glance duration (TGT) and the number of items is linear, indicating that no such anticipation is possible for unstructured interfaces (as would be expected).

$$TGT_{hh\_un} = 0.049N - 0.091 \; (R^2 = 0.998) \tag{13}$$

Relationships for NG can be derived in a similar fashion.

$$NG_{hh\_st} = 1 \tag{14}$$

$$NG_{hh\_un} = 0.0071N + 0.96 \ (R^2 = 0.98) \tag{15}$$

3.5. Refining the Model

It was evident from the data that, given the short duration of each 'task' (i.e. selecting a single item), experienced users were, on average, able to achieve selections (including pointing, decision and search activities) from a structured array in one glance, even for the larger target arrays (Equation 14). This suggests that *pointing* seldom necessitated a separate, dedicated glance as initially assumed. Instead, it was 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 behaviour, 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 constituted only a proportion of the final glance. Thus, we concluded that the final glance constituted elements of both *search/decide* and *pointing*. This is perhaps unsurprising given that in a divided attention context, it is expected that users would naturally be inclined to select a target as soon as it is located, rather than actively segregating decision/search and pointing activities (e.g. by returning their attention to the primary task), as this would require them to relocate the target during subsequent glances.

We therefore modified our analysis approach, disregarding a separate *pointing* element for NG, and assume that this glance is already included during the derivation of NG_{ds}. Again, we can derive *MGD* by dividing *TGT* (i.e. TGT_{pt} plus TGT_{ds}) by *NG*, in line with International Standards definitions [14]. Combining *searching*, *decision* and *pointing* terms, the following equations were therefore obtained:

$$TGT_{st} = \left(\frac{\log_2 N}{\log_2(N+t_i)}\right)(0.029N + 0.44) + 0.11\log_2 N + 0.11\log_2\frac{D}{W} + 0.35 \tag{16}$$

$$TGT_{un} = \left(\frac{\log_2 N}{\log_2(N+t_i)}\right)(0.10N - 0.028) + 0.049N + 0.045\log_2 N + 0.11\log_2\frac{D}{W} + 0.17 \tag{17}$$

$$NG_{st} = \left(\frac{\log_2 N}{\log_2(N+t_i)}\right)(0.021N + 1.04) + 1 \quad (18)$$

$$NG_{un} = \left(\frac{\log_2 N}{\log_2(N+t_i)}\right)(0.044N + 0.81) + 0.0071N + 1.96 \quad (19)$$

$$MGD_{st} = TGT_{st} \div NG_{st} \quad (20)$$

$$MGD_{un} = TGT_{un} \div NG_{un} \quad (21)$$

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

3.6. Study 3: Model Validation

To ensure thorough evaluation, results obtained for TGT, NG and MGD (obtained from testing a novel touchscreen interface in accordance with the NHTSA EGDS test protocol [1]) were initially compared to predictions made by both *structured* and *unstructured* predictive models (Equations 16-21). 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.

Moreover, it was assumed that each subtask (and its associated metrics) occurred in series and thus predictions could be aggregated to determine the total (TGT, NG) or mean (MGD) visual demand associated with each complete task. Figure 7 shows observed behaviour plotted against the structured and unstructured predictions for each task.

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.

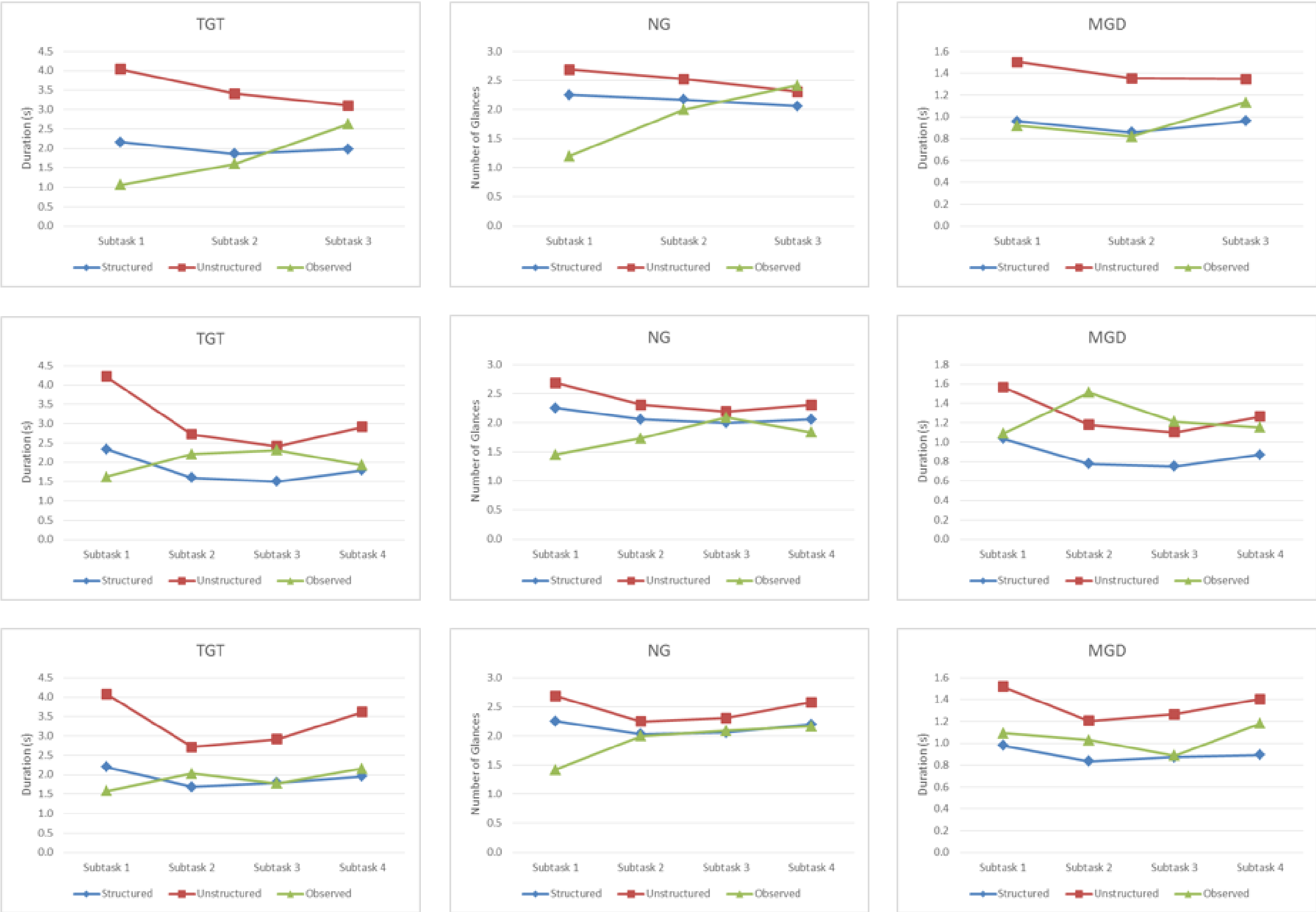


Fig. 7. *Structured and unstructured model predictions compared to observed performance, isolated by subtask, for Entertainment (top), HVAC (centre) and Personal Comfort (bottom) tasks. Predictions are shown in blue (structured) and red (unstructured) lines; observed data is shown in green.*

4. Discussion

We describe a model of human behaviour inspired by the work of Cockburn et al. [12] that aims to predict the demand of finger-touch ‘search, find and select’ tasks in a driving context, based on information theory principles (Fitts’ Law and Hick-Hyman Law). The model therefore extends the work of Cockburn et al [12] and in particular makes this applicable to visual demand in situations of divided attention, such as driving.

There are clear similarities between our model and the equations proposed by Cockburn et al. [12]. A notable difference is the inclusion of an additional visual search/experience scale factor (d_{vs}), which reflects the increased learning required for interfaces containing a larger number of targets, and the additional visual demand necessitated by switching between primary and secondary task execution. This is an important consideration because it reflects the fact that 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 the initial exposure to the interface. Consequently, an element of ‘searching’ is likely to persist, even after multiple interactions with well-designed, ‘learnable’ interfaces. In contrast, no such ‘chunking’ and search resumption would be expected in situations involving static menus in a sedentary context (where the interface is the primary – and often only – candidate for users’ attention); in these situations, the ‘searching’ element quickly decays during repeat exposure, as shown by Cockburn et al. [12].

An important consideration when designing HMIs 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 icon-based interface to be ‘unstructured’ when encountered for the first time. However, 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 interface as ‘structured’. By deriving equations for both structured and unstructured interfaces in the current study, it is therefore possible to predict the range of performance that may be achieved from initial exposure – typical of novice or occasional users – to expert performance, achieved through regular, repeated use.

It is evident from the preliminary validation study that, with the exception of the home-screen (subtask 1), the empirical data generally existed between structured and unstructured predictions, with notable variability for different subtasks/measures (see Figure 7), suggesting that this interpretation may be true – in situations where the observed behaviour was more closely aligned with ‘unstructured’ predictions, users were less familiar with that particular subtask or screen layout, and one might consequently expect visual behaviour to migrate towards ‘structured’ predictions as drivers’ 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, suggesting that this was very familiar and easy to access.

An alternative explanation is that the observed behaviour may be indicative of the effectiveness of the interface design. Indeed, participants who took part in the validation study were trained in accordance with NHTSA EGDS protocol [1] and thus, could be considered as ‘expert performers’. Consequently, 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 effective and afforded anticipation. In contrast, visual behaviour tending towards the unstructured predictions may be indicative of poor design, i.e. the interface or task lacked learnability. Therefore, the predictive model may also have utility as a formative design evaluation tool. However, it is noteworthy that the evaluation results were predicated on the assumption that the tasks could be broken down into discrete subtasks, each of which could be modelled independently and occurred in series; 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 design techniques, such as skeuomorphic elements that may not lend themselves to theoretical analysis. 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 the 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 (see: [15]), and thus the derived equations would have differed. Visual demand may also have been influenced by the search tasks employed to capture data, such as semantic complexity associated with different target words in study 2, although efforts were made to mitigate this affect by selecting unambiguous targets. Furthermore, in a real-world environment, other factors that are likely to vary significantly between vehicle designs (e.g. space, anthropometry, location of touchscreen, the provision of an arm rest to support touchscreen operation, drivers' handedness 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. Consequently, using the model to provide *absolute* predictions of visual demand should be considered carefully and we recommend that derived data should serve as a guide only. The model's true utility is likely to exist in its ability to predict *relative* metrics, thereby allowing several prospective design concepts to be considered early in the design cycle, and reducing arguments to simple calculations based on an understanding of the underlying characteristics of the interface and task (see: [2]).

5. Conclusions and Future Work

We describe a predictive model of visual behaviour associated with in-vehicle HMIs that combines elements from Fitts' Law and Hick-Hyman Law and considers anticipation afforded by structuring and repeated exposure to an interface; it also refelcts the additional learning required to achieve expert status while using menus containing larger numbers of items. Initial validation suggests that the model may be effective as both an evaluation and design tool, enabling stakeholders to consider the visual demand associated with a larger number of HMIs or menu designs, intended for in-vehicle deployment, much earlier within the design cycle and without the pre-requisite investment in costly implementation or extensive user trials. In line with similar theoretical work (e.g. [12]), the model assumes that 'search, find and select' tasks can be considered as comprising separate decision/search and pointing components that can be modelled independently; this may be an over-simplification and further validation work is required to validate this approach. 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 and skeuomorphic elements, intended to enhance HMI usability, learning and design.

6. Acknowledgements

The research was conducted in collaboration with Jaguar Land Rover Research and the authors would like to gratefully acknowledge their support.

7. References

[1] National Highway Traffic Safety Administration, "Visual-Manual NHTSA Driver Distraction Guidelines For In-Vehicle Electronic Devices," 2013.

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

[3] M. Pettitt, G. Burnett and A. Stevens, "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*, 2007.

[4] C. E. Shannon and W. Weaver, "The mathematical theory of information," 1949.

[5] S. K. Card, W. K. English and B. J. Burr, "Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT," *Ergonomics,* vol. 21, no. 8, pp. 601-613, 1978.

[6] P. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *Journal of experimental psychology,* vol. 47, no. 6, p. 381, 1954.

[7] W. Hick, "On the rate of gain of information," *Quarterly Journal of Experimental Psychology,* vol. 4, no. 1, pp. 11-26, 1952.

[8] R. Hyman, "Stimulus information as a determinant of reaction time," *Journal of experimental psychology,* vol. 45, no. 3, p. 188, 1953.

[9] W. Beggs, J. Graham, T. Monk, M. Shaw and C. Howarth, “Can Hick's law and Fitts' law be combined?,” *Acta psychologica,* vol. 36, no. 5, pp. 348-357, 1972.

[10] R. Soukoreff and I. MacKenzie, “Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts’ law research in HCI,” *International journal of human-computer studies,* vol. 61, no. 6, pp. 751-789, 2004.

[11] E. Hoffmann and J. Lim, “Concurrent manual-decision tasks,” *Ergonomics,* vol. 40, no. 3, pp. 293-318, 1997.

[12] A. Cockburn, C. Gutwin and S. Greenburg, “A predictive model of menu performance,” in *Proceedings of SIGCHI conference,* NY, 2007.

[13] S. G. Klauer, T. A. Dingus, V. L. Neale, J. D. Sudweeks and D. J. Ramsey, “The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data,” 2006.

[14] International Organization for Standardization, “Road vehicles – Measurement of driver visual behaviour with respect to transport information and control systems – Part 1: Definitions and parameters,” 2014.

[15] D. Large, E. van Loon, G. Burnett and S. Pournami, “Applying NHTSA task acceptance criteria to different simulated driving scenarios,” in *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications,* Nottingham, 2015.