This is the post-print (after peer-review) version of the paper which appeared in IET Intelligent Transport Systems, 6(3): 243-258 (not typeset)

The Trade-Off Between Context and Objectivity in an Analytic Approach to the Evaluation of In-Vehicle Interfaces

Catherine Harvey[a], Neville A. Stanton[a]

[a] Transportation Research Group, Faculty of Engineering and Environment, University of Southampton, Highfield Campus, Southampton, Hampshire, SO17 1BJ, UK

This paper presents a case study to explore an analytic approach to the evaluation of In-Vehicle Information Systems (IVIS) usability, aimed at an early stage in product development with low demand on resources. Five methods were selected: Hierarchical Task Analysis (HTA), Multimodal Critical Path Analysis (MCPA), Systematic Human Error Reduction and Prediction Approach (SHERPA), Heuristic Analysis, and Layout Analysis. The methods were applied in an evaluation to two IVIS interfaces: a touch screen and a remote controller. The findings showed that there was a trade-off between the objectivity of a method and consideration of the context of use: this has implications for the usefulness of analytic evaluation. An extension to the Multimodal Critical Path Analysis (MCPA) method is proposed as a solution to enable more objective comparisons of IVIS, whilst accounting for context in terms of the dual-task driving environment.

1. Introduction

This case study explores the use of analytic modelling in the In-Vehicle Information System (IVIS) development cycle. The motivation for the work was to understand how to deliver an approach to modelling aspects of IVIS usability, working with inevitable commercial constraints, to provide useful information on which to base design decisions. IVIS have been used by automotive manufacturers, particularly in the premium vehicle sector, for the last decade. These systems integrate many secondary vehicle systems, including entertainment, comfort, communications and navigation, into a single screen-based interface. IVIS are most commonly controlled by the driver via a touch screen, remote controller, hard buttons, or a combination of these devices. Recently the issue of IVIS usability has received growing attention [1-3]. This is commensurate with the increase in functionality of these systems [4-6], which has been accompanied by the introduction of new approaches to facilitate the user-system interaction [7, 8]. The issue of usability of such interfaces is now more significant than ever.

1.1 Defining Usability

An essential starting point for assessing usability is to define criteria against which to evaluate. Since terms such as 'ease of use' [9], 'user friendliness' [10] and 'user perceived quality' [11] were first introduced, there have been many definitions of the concept of 'usability': in particular see Shackel [12], Norman [13], Nielsen [14], Shneiderman [15], Bevan [16] and The International Organization for Standardization [17]. One of the most widely used definitions is found in ISO 9241, part 11: guidance on usability [17], which defines usability as:

> [The] extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. [17, p.2]

The reason for the wide adoption of this definition is probably the reference to 'context of use'. Despite many attempts to produce a universal definition of usability, the importance of the context of use in defining specific usability criteria means that definitions need to be constructed according to characteristics of individual products and the environments in which they are used [6].

1.2 Usability evaluation

Harvey et al., [18] proposed a framework for the evaluation of IVIS, as shown in figure 1. The framework follows a mixed-methodology approach, incorporating analytic and empirical methods to model different aspects of usability. Analytic methods can be used to evaluate IVIS without needing to simulate the interaction with real users. This evaluation involves the examination of the intrinsic features of a product or task, i.e. task structure, and results in predictions about performance, including task interaction times, error rates, usability issues, and interface design [19]. Analytic methods require data about the tasks, users and system which comes from paper-based specifications and expert knowledge. In contrast, empirical evaluation methods measure performance directly [19], using existing or prototype systems, under simulated or real world conditions. This study focusses on the application of analytic methods only.

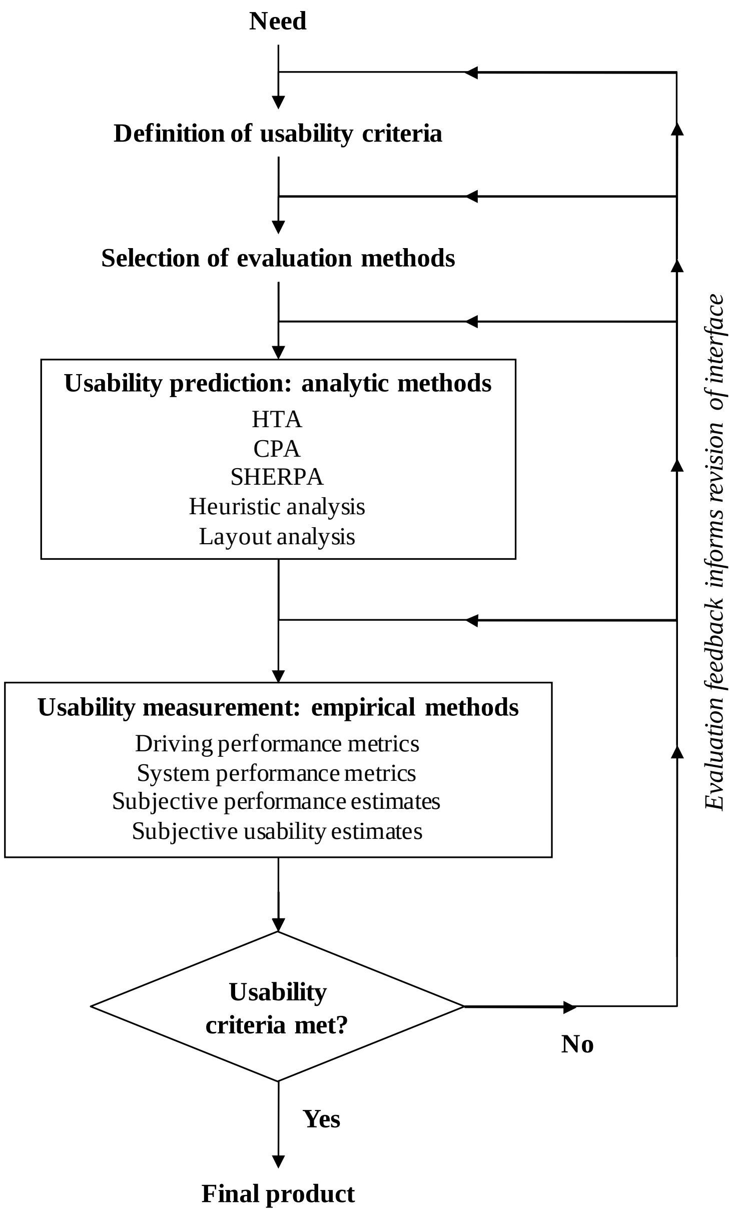

Figure 1. Framework for usability evaluation, adapted from Harvey et al. [18].

The aim of this study was to explore the usefulness of analytic methods for interface usability evaluation in the context of IVIS. Analytic methods were selected to meet a requirement for an approach to evaluation which can be applied at an early stage of product development with little demand for resources; however, currently these methods are not widely used in the automotive industry for IVIS evaluation. This study therefore attempts to explore the utility of analytic methods, including advantages and disadvantages, identify training and application times, and address shortcomings by proposing the extension to one or more of the techniques to increase their utility in a driving context. The findings will be useful to interface designers and evaluators working within the automotive industry, but also in other domains, to support the selection and application of analytic methods, with the overall objective of encouraging early-stage evaluation and design for usability.

The study evaluated two IVIS interfaces: a touch screen, which is one of the most commonly used interface types; and a remote controller, which works like a joystick to control a cursor on screen and was recently introduced to the market. It is important for automotive manufacturers to evaluate the performance of a new IVIS interface technology like the remote controller against their current system, as a benchmarking activity. The results of this comparison are reported in the case study; however, the main aim was to explore the intrinsic attributes of analytic methods in the context of IVIS interface evaluation [19], rather than as a direct comparison of systems.

1.3 Analytic Methods

Analytic methods were selected to model the different aspects of IVIS interface performance: task structure, interaction times, error rates, usability issues and interface design. Principles to guide the selection of evaluation methods were defined in a previous study [18]: consider the type of information required from the evaluation, consider the stage of the design process at which evaluation is needed, consider the resources required and those available, and consider the people that will need to be involved in the evaluation. Today, usability evaluation is encouraged in academia and industry; however, there have been suggestions that it can be ineffective and even detrimental if applied blindly and according to rule, rather than as a method of encouraging thought and consideration in designers and developers [20]. Automotive manufacturers also tend to employ two distinct approaches to IVIS evaluation: driving performance measures in relation to safety of driving whilst using an IVIS, and customer satisfaction measured by surveys [21]. The analytic methods presented in this study were selected to meet a requirement for measures which give an indication of interface usability before a product is sent to market and which encourage designers to explore how the design of an interface influences the user experience. A review of analytic methods was conducted and the five presented in this study were identified as most suitable in an IVIS context, given the constraints of the automotive industry described above. For a more detailed discussion of method selection see Harvey et al. [18]. Descriptions of the five methods and the related IVIS performance factors are presented in table 1. Table 2 lists the inputs and outputs of each method. Heuristic Analysis and Layout Analysis yield mainly qualitative data; Multimodal Critical Path Analysis (MCPA) and Hierarchical Task Analysis (HTA) are used to generate mainly quantitative data; and Systematic Human Error Reduction and Prediction Approach (SHERPA) produces both quantitative (error rate) and qualitative (remedial strategies) data [22]. Both types of information are discussed in relation to all five methods in order to explore the most useful applications of each method, e.g. for making direct comparisons or generating design recommendations.

Table 1. Analytic methods and related factors of IVIS interface performance.

Factors	Analytical methods	Description
Task structure	Hierarchical Task Analysis (HTA)	A task description method, used to break down tasks into their smallest components and structure them in a hierarchy of goals, sub-goals and plans [23, 24]. Although HTA normally needs to be combined with other techniques to produce meaningful analysis [25], it can illustrate where tasks might lead to ineffective interactions due to poor structure.
Interaction times	Multimodal Critical Path Analysis (MCPA)	Used to model task times based on the interactions between operations performed in different modes [26-28]. MCPA was selected over other time-prediction methods as it enables operations to be modelled in parallel. Task times produce high correlations with eyes-off road time [29, 30], which is a measure of the interference of secondary tasks in the dual task environment.
Error rates	Systematic Human Error Reduction and Prediction Approach (SHERPA)	Predicts error rates and types for particular systems and tasks [31, 32]. Errors will be useful in assessing the level of training which is needed for successful use of a product or system. The nature of the dual task driving environment will also give rise to specific errors, such as failing to complete an operation due to a sudden increase in primary task demand.
Usability issues	Heuristic Analysis	Uses a checklist of principles as a guide for identifying usability issues with an interface [14]. The content of the analysis is set according to the criteria of interest: dual task environment, environmental conditions, range of users and training provision. Because it is a subjective technique, it is less easy to predict factors such as uptake, which needs to be evaluated with real users.
Interface design	Layout Analysis	A method for evaluating an interface based on the position of related functions, according to frequency, sequence and importance of use [24]. It is related to the dual task criterion because the location of an IVIS in relation to the driver will affect the optimisation of layout. It is also related to frequency of use because familiarity of users with the interface is a factor which determines layout.

Table 2. Inputs and outputs for analytic methods.

	Inputs	Quantitative outputs	Qualitative outputs
HTA	Task specification	Number of operations, hierarchical task structure	Understanding of task, goals and plans
MCPA	HTA	Task interaction times	Operation dependencies
SHERPA	HTA	Error types and frequencies	Remedial strategies
Heuristic Analysis	Experience of system / task specification	Number of usability issues identified	Types of usability issues, potential problems
Layout Analysis	System layout diagrams	Number of layout changes required	Changes to interface layout

2. Method

An evaluation of two existing IVIS was performed using the five analytic methods in order to explore the utility of this approach, in terms of information inputs and outputs, training times, resource demands, and possible extensions, in the context of early-stage product development. The interfaces under investigation were a touch screen and a remote controller input device. Figure 2 illustrates the typical layout of these interfaces in a right-hand drive vehicle, showing the position of the display screen and additional control pad (this was only present in the remote controller system). The schematic shows the approximate positions of the interface features and is not to scale. The control pad had a similar function to a joystick, moving a pointer on screen.

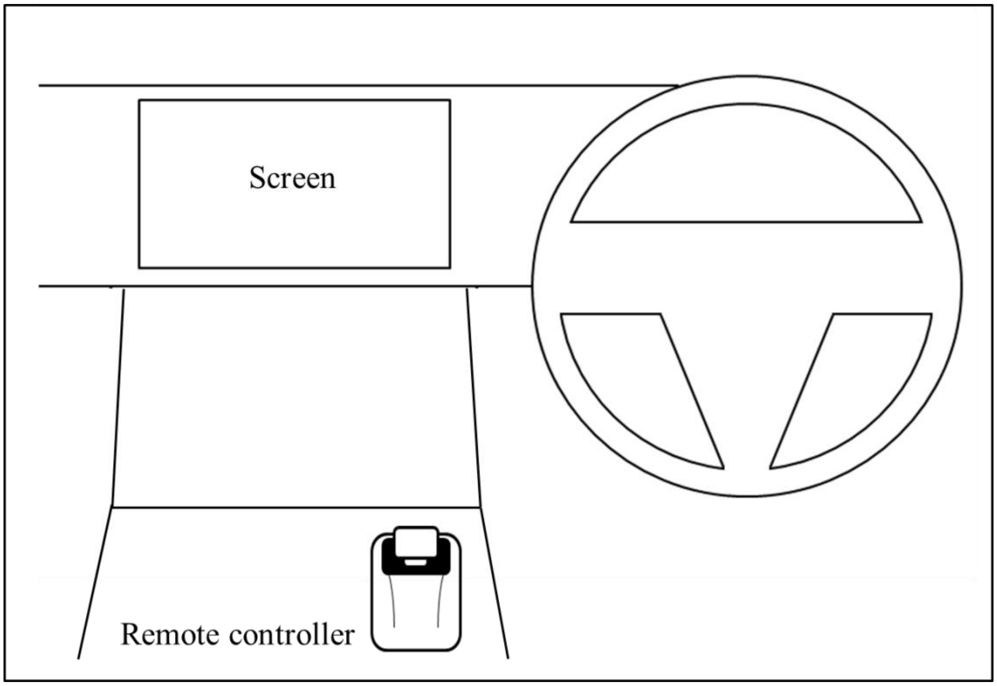

Figure 2. IVIS schematic.

2.1 Procedure

Expert walkthroughs of two existing IVIS were performed by a Human Factors analyst. These were based around a scenario of interacting with several in-vehicle, secondary tasks in a stationary vehicle. The term 'task' is used to refer to a sequence of operations performed by a user to achieve a goal, such as selecting a radio station or reducing fan speed. A single analysts applied all five methods reported in this study: Heuristic Analysis was performed first, whilst the analyst was interacting with each interface; the other four methods were applied after the data collection phase using the information gathered from each IVIS, in the order HTA, MCPA, SHERPA, Layout Analysis. The analyst trained in each of the methods prior to the data collection phase and spent approximately 4-5 hours using the two IVIS interfaces before applying the methods. A set of nine typical IVIS tasks was defined for this study as shown in table 3:

Table 3. IVIS tasks analysed in evaluation.

Categories	Tasks
Audio	Play radio station: 909AM (radio is currently set to 97.9FM)
	Increase bass by two steps
Climate	Increase temperature by 1°C (via centre console controls, not IVIS)
	Reduce fan speed by two steps
	Direct air to face only (air direction is currently set to face and feet)
	Direct air to face and feet (air direction is currently set to windscreen only)
	Activate auto climate (via centre console controls, not IVIS)
Navigation	Set navigation destination from system memory: ‘Home’
	Set navigation destination from previous destinations: ‘University Road, Southampton’

These nine tasks were selected from a set of over 130 tasks which were identified for existing IVIS from a review of automotive manufacturers’ IVIS manuals, which was conducted by the analyst prior to data collection. Four factors, defined by Nowakowski and Green [33], were used to guide task selection: use whilst driving, availability in existing systems, frequency of interaction and compliance with the 15 second task time rule [see 29]. The nine tasks were all likely to be used whilst driving, unlike other functions such as vehicle or display settings. Based on information from automotive manufacturers and the analyst’s personal experience, it was expected that the tasks would all be used at least once during a typical medium-long journey. All of the tasks were available in existing IVIS, including the two systems under investigation. Finally, preliminary investigations conducted by the analyst indicated that it should be possible to complete each of the nine tasks in less than 15 seconds. The 15 second rule, which is commonly referred to in the design and evaluation of IVIS tasks [29, 33, 34], states that no navigation tasks involving a visual display and manual controls, and available during driving, should exceed 15 seconds in duration [29]. Tasks were performed using each system and the inputs (from user to system) and outputs (from system to user) were recorded. Pictures were taken of the IVIS menu screens and controls at each stage of the interaction and the analysts recorded a description of each interaction. For the Heuristic Analysis, each IVIS was assessed against a checklist, developed by Stevens et al. [35]. The checklist was adapted for this evaluation by removing sections which were not directly connected to usability, including those relating to the documentation supplied with an IVIS, the packaging of the product, compliance with traffic regulations, system maintenance, and information referring to the road network.

2.2 Data Analysis

The data collected on each IVIS were modelled using the five analytical evaluation methods described previously. During the modelling phase, close attention was paid to the utility of each method and to the training times, execution times and resources required. Each of the methods and their application in this particular context of use is described further in the following sections.

3. Results and discussion

3.1 Hierarchical Task Analysis (HTA)

HTA was conducted for the two IVIS under investigation and an example of a HTA for the remote controller IVIS task ‘play radio station’ is presented in figure 3.

```
1 Play radio station
Plan 1 - Do 1-4 in order; WHEN radio station is set THEN 5

     1.1 Move hand from steering wheel to controller
     1.2 Open audio menu
     Plan 1.2 - Do 1, 2 in order

          1.2.1 Move pointer to AUDIO symbol
          1.2.2 Press ENTER button

     1.3 Select source from FM, AM or DAB
     Plan 1.3 - For FM THEN 1; for AM THEN 2, for DAB THEN 3

          1.3.1 Open FM menu
          Plan 1.3.1 - Do 1, 2 in order

               1.3.1.1 Move pointer to FM tab
               1.3.1.2 Press enter button

          1.3.2 Open AM menu
          Plan 1.3.2 - Do 1, 2 in order

               1.3.2.1 Move pointer to AM tab
               1.3.2.2 Press enter button

          1.3.3 Open DAB menu
          Plan 1.3.3 - Do 1, 2 in order

               1.3.3.1 Move pointer to DAB tab
               1.3.3.2 Press enter button

     1.4 Check for station in list
     Plan 1.4 - IF station is one of six presets THEN 1; IF not THEN 2

          1.4.1 Select radio station preset
          Plan 1.4.1 - Do 1, 2 in order

               1.4.1.1 Move pointer to radio station preset button
               1.4.1.2 Press ENTER button

          1.4.2 Search for radio station from frequencies list
          Plan 1.4.2 - Do 1, 2, 3 in order

               1.4.2.1 Move pointer to STATION LIST
               1.4.2.2 Press enter button
               1.4.2.3 Find and select station from list
               Plan 1.4.2.3 - To search through lower frequencies
               THEN 1; to search through higher frequencies THEN 2;
               do 3 to select station

                    1.4.2.3.1 Scroll up through radio station
                    frequencies
```

Figure 3. Excerpt of HTA for ‘play radio station’ task using remote controller IVIS.

3.1.1 IVIS evaluation

HTA is a task description method. Task description is a necessary precursor for further analysis, such as MCPA, which will produce measurable results [22]. HTAs for two or more systems may be subjectively compared in order to identify differences in task structure; however, this exercise is useful for task design exploration, rather than as a method for contrasting products. It is possible to compare two or more different products or individual tasks based on the number of operations identified by HTA. A system which requires the user to perform a large number of operations per task is likely to be less efficient than a system with fewer operations; however, this will also depend on the time taken to perform each operation and the error potential of the tasks involved.

3.1.2 Utility of the method

HTA is a fairly time consuming method to carry out as each individual operation in a task needs to be analysed; however, creating a comprehensive HTA can considerably reduce the time required for other modelling methods such as MCPA and SHERPA. A problem facing Human-Computer Interaction (HCI) is that interfaces are often engineering-focussed and are therefore not optimised for activity patterns [36]. HTA provides an activity-based classification of user behaviour, which in itself can be used to improve interface design. The process of conducting HTA can also provide the analyst with important information about task structure and menu design. A deeper understanding of the links between task design and usability should increase focus on good HMI design: we therefore also recommend that designers and Human Factors specialists within manufacturing companies use the process as a learning tool.

3.2 Multimodal Critical Path Analysis (MCPA)

HTA is used as a starting point for MCPA. In this study, operations identified in the HTA were further categorised as visual, manual, cognitive or auditory. MCPA was selected here in preference to other theoretical modelling methods because it is the only one capable of modelling parallel activities in multiple modes [26, 37, 38]. Parallel activities can be described according to the multiple resource model

[39], which proposes that attention can be time-shared more effectively between operations across different modes, compared with operations which utilise the same mode [39]. Two visual operations, for example locating a control on the vehicle's dashboard and reading a label on screen, cannot occur in parallel; however, one of these visual operations may take place at the same time as a physical operation, such as moving the hand towards the screen. The structure of a MCPA model is also affected by the dependency of operations. A dependent operation is one which cannot begin until a previous operation has been completed [26, 40]. For example, the driver cannot press the enter key on the remote controller until the pointer has been moved to the on-screen target. Figure 4 shows an extract from a MCPA diagram for the touch screen IVIS task: play radio station. Each operation is represented pictorially as a node and the relationships between operations are denoted by their relative positions in the MCPA diagram and by the arrows connecting each operation [28]. In figure 4, the operations 'locate audio/TV icon' (visual), 'move hand to touch screen' (manual) and 'make selection' (cognitive) are performed in parallel and are therefore presented one above the other. The operation 'homing on target' (manual) is dependent on the user having already decided on the correct button, located it on screen and moved their hand to the touch screen: it is therefore presented as a sequential operation in the diagram. The operation 'touch audio/TV button' (manual) is dependent on the user having homed in on the target with their finger; therefore this is presented to the right of the preceding 'homing' operation. MCPA diagrams were generated in this way until all operations in each task had been represented.

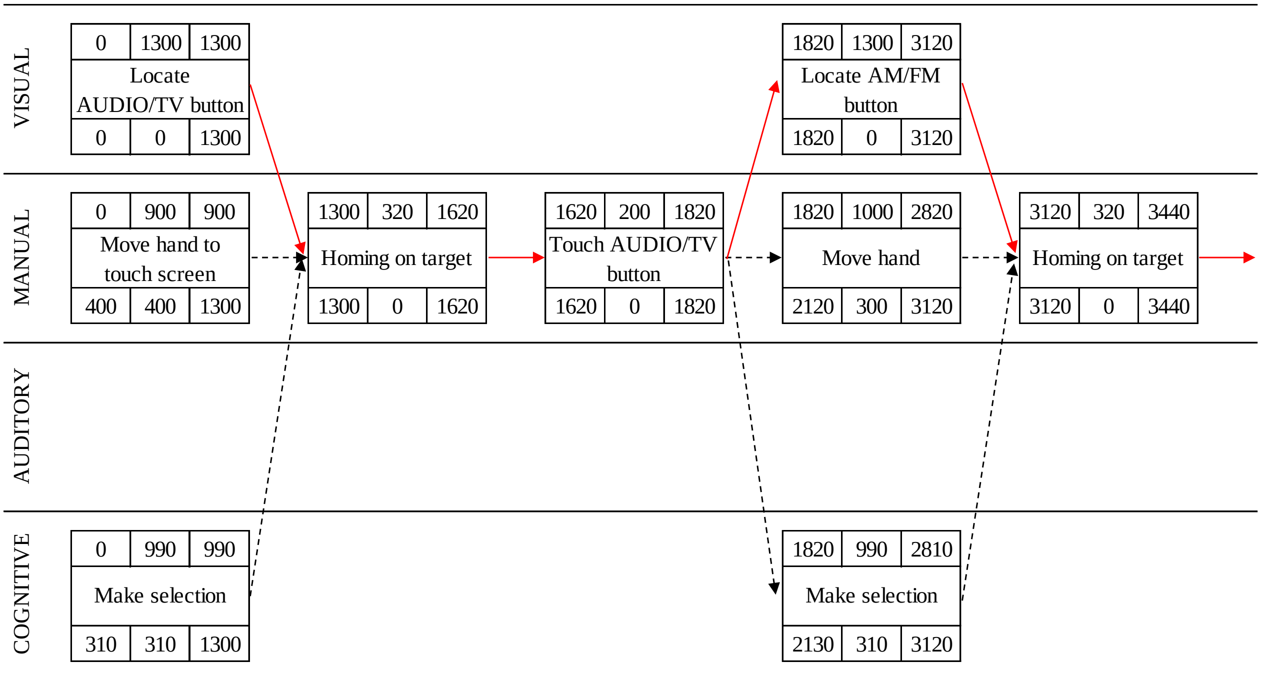

Figure 4. Excerpt of MCPA diagram for touch screen audio task, with operation timings given in ms.

In the MCPA diagram, time flows from left to right; therefore, a succeeding operation which is dependent on the completion of a preceding operation is positioned to the right of the operation upon which it is dependent. Parallel operations are located in the same vertical position in the diagram and are separated into rows to represent the different interaction modes (visual, manual, cognitive, auditory). After modalities and dependencies are defined, durations can be assigned to each operation. In this study, these operation duration times were derived from a review of the HCI literature and are listed, along with their sources, in Table 4.

Table 4. Operation timings from HCI literature.

Mode	Task	Time (ms)	Reference	Time used in model
Visual	Locate target on screen	1300-3600	Stanton and Baber (2008)	Visually locate single target: 1300
	Recognise familiar words or objects	314-340	Olsen and Olsen (1990)	Visually locate sequential alphanumeric target: 340
	Check if on-screen target is highlighted	600-1200	Pickering et al. (2007)	Check if target is highlighted: 900
	Read page of text on screen, e.g. navigation warning	5000	*Average from performing task*	Read navigation warning: 5000
	Read number (e.g. temperature) on centre console display	1000-1200	Wierwille (1993)	Check temperature display: 1000
Manual	Move hand from steering wheel to touch screen/remote touch controller, and vice versa	900	Mourant et al. (1980)	Move hand to touch screen or remote controller: 900
	Press button/target	200	Baber and Mellor (2001)	Press touch screen target: 200
	Move hand between targets	400	Card et al. (1983)	Homing on target (movement time during visual search assumed extra): 320 [total 520 with touch target time]
		505-583	Ackerman and Cianciolo (1999)	
		520	Stanton and Baber (2008)	
		368-512	Rogers et al. (2005)	
	Move pointer to target on screen	1290	Card et al. (1978)	1290 includes pressing enter; therefore, positioning time: 1290-570=720
	Press hard enter button	570	Card et al. (1983)	Press enter button on remote controller: 570
				Press button on centre console: 570
Auditory	Listen for feedback to confirm correct radio station	3000	*Average from performing task*	Listen for radio station confirmation: 3000
	Listen for change in audio settings (e.g. bass)	3000	*Average from performing task*	Listen for audio settings confirmation: 3000
Cognitive	Make simple selection	990	Stanton and Baber (2008)	Make selection: 990

Where a range of values has been reported, an operation time was estimated from within that range based on the analyst’s experience and knowledge of the IVIS tasks. There are also a number of rules and assumptions which support the use of these timings in the MCPA models:

- Time to visually locate a target is 1300ms, following [37], for any single target and the first alphanumeric target in a sequence.
- Time to visually locate a target is 340ms for any sequential alphanumeric target after the first target in a sequence. It is assumed that users would be more familiar with the layout of an alphanumeric keyboard than with the other menu screens in each system, therefore search time for alphanumeric targets was reduced to 340 ms, following the time to recognize familiar objects reported by [41].
- No cognitive 'make selection' operation occurs in parallel with a sequential alphanumeric visual search (340ms), following the heuristics for Mental operators devised by Card et al. [42]. Entering a word into the system is assumed to be a single 'chunk': users make a decision about the sequence of letters or numbers at the start of the chunk, therefore individual decisions for each subsequent alphanumeric entry are assumed to be unnecessary.
- There is always some movement of the hand/fingers (touch screen) or the cursor (remote controller) during visual search. This movement follows the direction of gaze so only a small 'homing' movement is needed after the target is found [41]. This movement time varies with the visual search time. It is assumed that the movement starts just after visual search begins, therefore a value of 1000ms has been assigned in the models.

Duration, modality and dependency information is used to calculate Early Start Time (EST) and Early Finish Time (EFT) as part of the forward pass through the network; Late Start Time (LST) and Late Finish Time (LFT) as part of the backward pass through the network; and finally, Float Time, according to the following rules:

The forward pass calculates the EST and EFT of each operation, moving progressively through the task diagram from left to right [28]. The EST of operation 'X' is determined by the EST of the preceding operation plus its duration. If there is more than one preceding operation which links into operation 'X', then the EST is determined by the latest EFT of the preceding activities:

EST of operation ‘X’ = EST of preceding operation + Duration of preceding operation

The EFT is the EST of an operation plus its duration time:

EFT of operation ‘X’ = EST of operation ‘X’ + Duration of operation ‘X’

The backward pass calculates the LST and LFT of each operation, starting from the ‘End’ node and moving from right to left, back through the task diagram. The LST of operation ‘X’ is determined by the LST of the succeeding activity minus the duration of operation ‘X’ [28]. If there is more than one succeeding operation that links directly into operation ‘X’ then the earliest possible LST should be used:

LST of operation ‘X’ = LST of succeeding operation – Duration of operation ‘X’

The LFT of an operation is determined by the sum of the LST and duration of an operation:

LFT of operation ‘X’ = LST of operation ‘X’ + Duration of operation ‘X’

After calculating the values from the forward and backward passes, the free time, or ‘float’, is calculated. All paths through the task network, with the exception of the critical path, will have some associated float time. Float time of operation ‘X’ is the difference between the LST and EST of operation ‘X’:

Float time of operation ‘X’ = LST of operation ‘X’ – EST of operation ‘X’

The final stage of MCPA involves defining the critical path and calculating total task time. The critical path occurs along the path of operations which has the most minimal float time: in figure 4, this is denoted by the solid lines. The durations of all operations on the critical path are summed to produce the total task time.

3.2.1 IVIS evaluation

Total task times were calculated for the touch screen and remote controller interfaces and are presented in table 5, along with the differences between the two devices for each task.

Table 5. Total task times for secondary tasks performed via the touch screen and remote controller IVIS.

Task	Task time (ms)		Difference	
	Touch screen	Remote controller	Remote controller - Touch screen	%
Play radio station (909AM)	8460	10770	2310	27.30
Increase bass by 2 steps	11380	11100	-280	-2.46
Increase temperature by 1 degree	4860	4860	0	0.00
Reduce fan speed by 2 steps	5340	6650	1310	24.53
Direct air to face and feet	8880	6080	-2800	-31.53
Direct air to face only	7060	6080	-980	-13.88
Turn on auto climate	3090	3090	0	0.00
Enter destination from system memory	16820	11260	-5560	-33.06
Enter destination from previous entries	16820	11260	-5560	-33.06
Total task time	82710	71150	-11560	-13.98

The MCPA method predicted that five tasks would take longer with the touch screen than the remote controller and that two tasks would take longer with the remote controller than the touch screen. There was no difference between the two systems for the ‘increase temperature’ and ‘auto climate’ task times: this was because they were performed via centre console controls rather than the IVIS interfaces and the task design was identical in both cases. The two air direction tasks were predicted to be shorter with the remote controller than the touch screen. In the remote controller system the user is allowed to select the exact options directly because there are separate options for air to ‘face and feet’ and ‘face only’; however, the touch screen presents three options (‘face’, ‘feet’, and ‘windscreen’) and the user therefore needs to select multiple options to set air direction to face and feet: this involves extra operations to complete the task. The destination entry tasks were also predicted to take longer with the touch screen compared to the remote controller. This is because the touch screen system required users to read a warning about using the navigation function whilst driving and this contributed a large amount of time to the task (5000ms to read the warning, 1300ms to locate the ‘Agree’ button, 320ms homing time to target, 200ms to touch target: 6820ms total extra time). Without this extra task segment, the touch screen would have produced shorter task times than the remote controller for the two navigation tasks. Similarly, the time difference in the 'increase bass' task can be attributed to an extra task segment in the touch screen task: with this system, the user had to select the 'Audio/TV' button, then 'Settings', followed by 'Sound', in order to access the 'Bass +' target; however, with the remote controller system, the 'Settings' menu is eliminated and the user moves directly from the 'Audio' menu to the 'Sound' screen. This analysis shows that the time differences between the two IVIS for the air direction, navigation and increase bass tasks resulted from differences in task design between the two systems, in other words, it is the extra steps involved in the touch screen tasks which were responsible for the observed differences in task times, rather than differences in the nature of the interface. The effect of interface design can, however, be seen in the predicted times for the other tasks including 'play radio station' and 'reduce fan speed'. MCPA predicted shorter times for these tasks with the touch screen, compared to the remote controller. When the individual task segments are examined, it appears that the nature of inputs to the touch screen system supports quicker performance because the individual operations have shorter durations, as illustrated in figure 5.

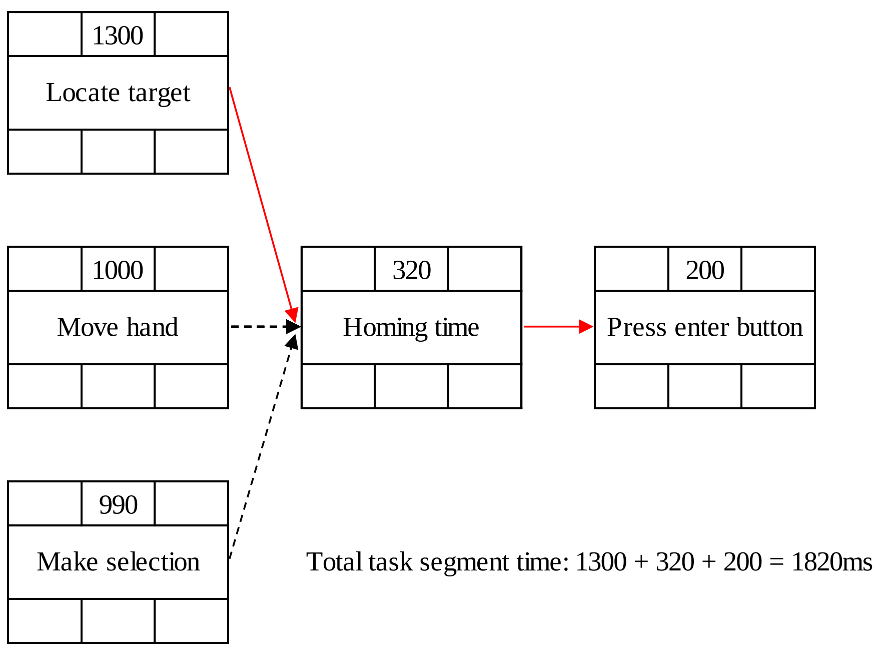

Figure 5. Excerpt from touch screen MCPA diagram to show a single target selection segment.

This can be compared to a task segment from a remote controller MCPA diagram, showing the same target selection activity, as illustrated in Figure 6.

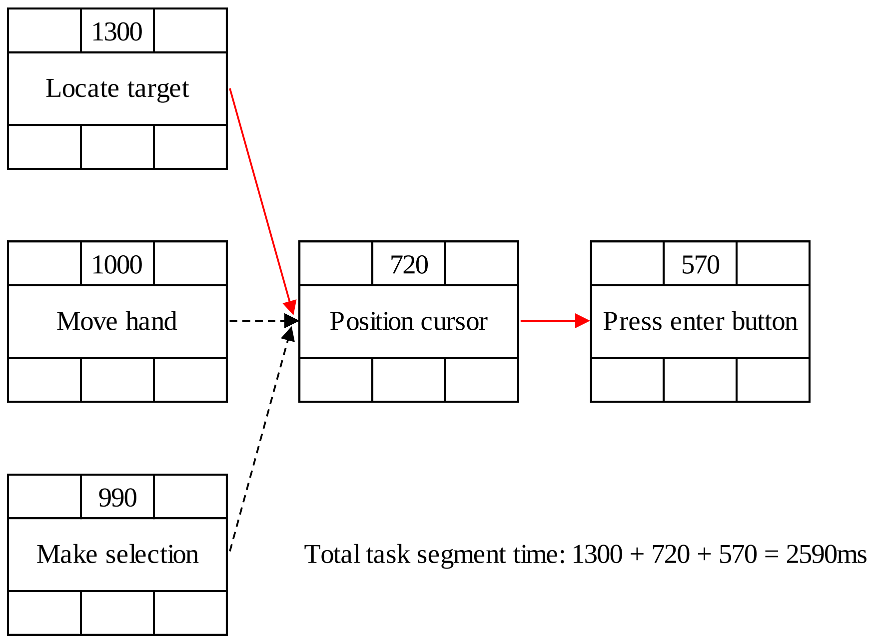

Figure 6. Excerpt from remote controller MCPA diagram to show a single target selection segment.

The location options are of equal duration for both IVIS as this operation requires the user to visually locate a target on screen and the target and screen sizes were approximately equivalent for the two systems. The difference in segment time is produced by the second and third operations in the sequence, which involve the user either homing their hand/fingers to the touch screen target and pressing the target, or manipulating the remote controller to move the cursor to an on-screen target and pressing the enter button on the side of the controller. The touch screen operation times were based on the time for moving the hand (320ms) and pressing a key (200ms) reported by Stanton and Baber [37] and there is also some assumed movement of the hand which occurs in parallel with the visual search operation. Previous studies have reported times of between 368ms and 583ms for physical selection of on-screen targets, combining movement of the hand and pressing a target [42-44], which are commensurate with those used in the current study (320 + 200 = 520ms). Card et al. [42] reported a time of 570ms for pressing a pushbutton and this value was used in the remote controller model for the time to press the enter button on the side of the controller. An assumption was made that pressing a hard enter key located on the side of the remote controller (570ms) would take longer than touching a target on screen (200ms) due to the increased resistance from the remote controller button and the reduced ease of access. Card et al. [45] reported positioning time for a mouse-controlled cursor as 1290ms, which included target selection via a button press. The movement of the remote controller was very similar to a mouse and it was assumed that this value provided a good approximation of positioning time for the remote controller. Time to press the enter button (570ms) was subtracted from total mouse positioning time (1290ms) to give a value of 720ms, which was assigned to the positioning of the remote controller in the model. This combination of positioning the cursor and pressing the enter button resulted in longer task segment times for the remote controller, compared with the touch screen, demonstrating that the nature of the interaction styles of the two devices were different and that this had an effect on total task times.

3.2.2 Utility of the method

MCPA enabled a quantitative comparison of task times to be made between the two IVIS, following a structured procedure based on information from the HTA. As this procedure was applied to both interfaces, it is likely that the relative comparisons had high construct validity. On the other hand, we cannot be sure that the results represent accurate measures of absolute task times, because they have not been validated against real interactions. There is potential for the MCPA method to model absolute task times accurately if a comprehensive and valid database of IVIS operation types could be developed. MCPA in its current form fails to address the issue of the dual task driving environment, as it does not account for breaks in task performance caused by the driver’s attention reverting back to the primary driving task. Although static task times have been found to correlate well with eyes-off-road time [29, 30], incorporating the split in visual attention into the model would produce more accurate predictions of IVIS task times in a dynamic environment. This study has highlighted further potential for developing the method to enable dynamic, dual-task environments to be accurately modelled. This will be investigated in future work.

3.3 Systematic Human Error Reduction and Prediction Approach (SHERPA)

SHERPA was applied to the two IVIS and operations were classified into one of five types: action, retrieval, checking, information communication, and selection [25]. This classification was based on the analyst's judgement. Within each error type there are a number of error modes, which are shown in table 6.

Table 6. Error modes and their descriptions

Error mode	Error description
Action	
A1	Operation too long/short
A2	Operation mistimed
A3	Operation in wrong direction
A4	Operation too much/little
A5	Misalign
A6	Right operation on wrong object
A7	Wrong operation on right object
A8	Operation omitted
A9	Operation incomplete
A10	Wrong operation on wrong object
Information retrieval	
R1	Information not obtained
R2	Wrong information obtained
R3	Information retrieval incomplete
Checking	
C1	Check omitted
C2	Check incomplete
C3	Right check on wrong object
C4	Wrong check on right object
C5	Check mistimed
C6	Wrong check on wrong object
Information communication	
I1	Information not communicated
I2	Wrong information communicated
I3	Information communication incomplete
Selection	
S1	Selection omitted
S2	Wrong selection made

Credible error modes were determined for each task step in the HTA along with the form that the error would take, the consequences of the error, and the recovery potential. Next, the analyst estimated the probability (*P*) of the error occurring during the task and also the criticality (*C*) of the error, using an ordinal scale: Low (L), Medium (M), High (H). Finally, the analyst proposed strategies to reduce the identified errors. An extract of a SHERPA output table for the touch screen IVIS task ‘play radio station’ is presented in table 7.

Table 7. Extract of SHERPA output table for touch screen IVIS task: play radio station

Task	Error mode	Error description	Consequence	Recovery	Probability	Criticality	Remedial strategy
1 Play radio	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.1 Open AUDIO/TV menu	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.1.1 Move hand to touch screen	A8	Driver cannot remove hand from wheel due to high primary task demand	Cannot perform any interaction with touch screen	Immediate, when primary demand allows	M	M	Reduce need for removing hands from wheel - increase number of steering wheel controls, increase automation of secondary tasks
	A9	Driver starts to move hand towards screen but has to replace on wheel due to sudden primary task demand	Cannot perform any interaction with touch screen	Immediate, when primary demand allows	M	H	Reduce need for removing hands from wheel - increase number of steering wheel controls, increase automation of secondary tasks
1.1.2 Prepare to open menu	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1.1.1.1 Make selection	S2	Wrong selection made	Incorrect menu opened	Immediate	L	M	Ensure labels clearly relate to function
1.1.1.2 Locate AUDIO/TV icon	R1	Visual check is not long enough to locate icon	Cannot open desired menu	Immediate, when primary demand allows	M	L	Make icons and labels larger to ensure quick identification
	R2	Incorrect icon is located by mistake	Wrong menu is opened if mistake is not realised	Immediate	L	M	Ensure icons clearly relate to function
1.1.3 Touch AUDIO/TV button	A4	System does not recognise touch	Audio/TV does not open	Immediate	H	L	Increase sensitivity of touch screen
	A6	Touch incorrect button or other part of screen	Incorrect input made or no input made	Immediate	M	M	Increase size of buttons

3.3.1 IVIS evaluation

SHERPA was performed for each of the nine tasks for both systems. Tables 8a and 8b present all identified errors and error modes for the touch screen and remote controller respectively. The tables also include the probability (*P*) and criticality (*C*) ratings for each error, shown in bold. Significant errors were defined as those with either high *P* or *C* ratings or where both *P* and *C* were rated as medium. The number of error descriptions, i.e. 'system does not recognize touch', which were rated as significant was used as a metric by which to compare the two IVIS interfaces. There was little difference in the number of significant error descriptions identified for the touch screen (6) and remote controller (7).

Table 8a. Errors identified by SHERPA for touch screen IVIS, including probability, criticality and frequency ratings.

Error mode	**Description**	***P***	***C***	**Frequency**
A2	Consecutive presses are too quick	M	L	3
A4	System does not recognise touch	**H**	**L**	24
A4	Press centre console button with too little force	M	L	2
A4	Repeat centre console button press too many times whilst waiting for accurate feedback	L	M	2
A5	User moves hand to wrong area of screen	M	L	17
A6	Touch incorrect button or other part of screen	**M**	**M**	24
A6	Touch incorrect button or other part of centre console	**H**	**M**	2
A8	Driver cannot remove hand from wheel due to high primary task demand	**M**	**M**	9
A8	Driver does not move hand back to steering wheel	L	L	9
A9	Driver starts to move hand towards screen but has to replace on wheel due to sudden primary task demand	**M**	**H**	9
A9	Operation incomplete, due to increased demand from primary task	**M**	**M**	17
R1	Visual check is not long enough to locate icon	M	L	26
R2	Incorrect icon is located by mistake	L	M	26
C1	Check omitted	L	L	9
C2	Check is not long enough to obtain accurate feedback	L	M	2
S2	Wrong selection made	L	M	26
			TOTAL ERRORS	**207**

Table 8b. Errors identified by SHERPA for remote controller IVIS, including probability, criticality and frequency ratings

Error mode	**Description**	***P***	***C***	**Frequency**
A4	Press button with too little force	L	M	22
A4	Repeat button press too many times whilst waiting for accurate feedback	L	M	2
A5	Pointer misses icon/button/letter/number	**H**	**L**	20
A6	Select incorrect icon/button/letter/number	**M**	**M**	20
A6	Press down controller instead of enter button located on side of controller	**H**	**L**	20
A6	Touch incorrect button or other part of centre console	**H**	**M**	2
A8	Driver cannot remove hand from wheel due to high primary task demand	**M**	**M**	9
A8	Driver does not move hand back to steering wheel	M	L	9
A9	Driver starts to move hand towards controller but has to replace on wheel due to sudden primary task demand	**M**	**H**	9
A9	Driver cannot locate controller after physical search	**L**	**H**	7
R1	Visual check is not long enough to locate icon	M	L	22
R2	Incorrect icon is located by mistake	L	M	22
C1	Check omitted	L	L	7
C2	Check is not long enough to obtain accurate feedback	L	M	1
S2	Wrong selection made	L	M	22
			TOTAL ERRORS	**194**

3.3.2 Utility of the method

Within each system, many of the same errors were identified for each task because the tasks consisted of similar steps. All of the errors identified for both systems would have been likely to show up from analysis of only one or two representative tasks, which would reduce analysis time considerably. This should be a consideration for future development of error analysis techniques in this context. SHERPA was useful for investigating IVIS interactions in a dual-task environment, i.e. performing IVIS tasks at the same time as driving. For example, instances of incomplete tasks and failure to start tasks were predicted for situations in which the demand from primary driving was high; however, SHERPA provided no way of estimating the severity of these errors or the frequency with which they may occur. Although SHERPA follows a fairly rigid structure for assigning errors, the suggestions for remedial strategies for addressing those errors are likely to differ between analysts. SHERPA would benefit from repeated analyses by different personnel on a small sample of representative tasks. A focus group scenario, comprising a mix of ergonomists, designers and engineers, would also be a useful addition to the method to generate more useful remedial strategies.

3.4 Heuristic Analysis

The Heuristic Analysis was applied by the analyst, using an adapted IVIS checklist originally developed by Stevens et al. [35]. The checklist was organised into nine sections covering integration of the system into the vehicle, input controls, auditory properties, visual properties of the display screen, visual information presentation, information comprehension, menu facilities, temporal information, and safety-related aspects of information presentation. The evaluation was based on the analyst's experience, gained from four-five hours of interaction with each system, in a stationary vehicle.

3.4.1 IVIS evaluation

Tables 9a and 9b list the issues identified via the Heuristic Analysis for the touch screen and remote controller IVIS. The issues were categorised by the evaluator as positive or negative and further categorised according to the estimated severity of each issue. Negative (major and minor) and positive (major and minor) issues were identified for both IVIS using the heuristic checklist. There were slightly more negative issues identified for the remote controller (8), compared to the touch screen (7); however, because the difference was so small and the analysis was purely subjective, it is not possible to use these values to make a valid, quantitative comparison between the two systems.

Table 9a. Issues identified by Heuristic Analysis of touch screen IVIS

Negatives		Positives	
Minor	Major	Minor	Major
Little use of colour coding. Climate menus are cluttered. Discomfort from holding arm outstretched when operating touch screen.	Glare and reflections on screen. Easy to activate wrong control, especially with small buttons, and those which are close together. No non-visual location of touch screen controls. Small delay in screen response to touch for some functions.	Auditory feedback for touch screen button presses. Audio volume adjustable for some functions. Pop-up screens indicate extra information. Activation of functions via hard controls is confirmed by a message on screen.	Text information is easy and quick to read and understand (no long words or cluttered buttons). Easy to go back between menu levels and to return to HOME menu.

Table 9b. Issues identified by Heuristic Analysis of remote controller IVIS

Negatives		Positives	
Minor	Major	Minor	Major
Menu exits after a relatively short time. Colour coding not helpful. Auditory feedback volume not adjustable for button presses. Cluttered appearance of navigation screen. Unclear labeling of buttons. Temperature units not display.	Back button located in top right corner of screen: not easy to access, and no hard back button. Relative complexity of navigation menus.	Useful audio feedback to indicate incorrect entries. Large text size. Sensible use of abbreviations.	Screen is recessed, protected from glare, and located in a natural viewing position. Hard button for home menu. Easy non-visual location of hard controls.

3.4.2 Utility of the method

The Heuristic Analysis generated qualitative data relating to positive and negative features of each IVIS according to the checklist [35]. There are a number of checklists and guidelines for IVIS design [for example, see 35, 46, 47-49]; however, no single set of criteria has been accepted as the industry standard. This reflects the difficulty in defining a set of heuristics which is capable of providing a comprehensive checklist for IVIS usability. One of the main problems with the method was the lack of information regarding the frequency with which particular usability problems would occur in everyday usage. A further limitation of the heuristic method is the requirement for a fully developed product or prototype in order to evaluate some aspects of usability. This includes the effect of glare on the IVIS display screen, which cannot be assessed without exposing the IVIS to particular environmental conditions. This is a constraint imposed by the design of many existing checklists for IVIS evaluation, of which [35] is an example; however, it is possible that heuristics could be aimed at an earlier stage in design, eliminating the need for high fidelity prototypes. For example, Nielsen's 'Ten Usability Heuristics' [50] encourage a more general approach to usability evaluation which could be applied in the very earliest stages of product development. Based on these limitations, it is proposed that Heuristic Analysis could be a useful tool for reminding designers about important usability issues [38, 51], rather than for making direct comparisons between interfaces. The technique has potential for further development by individual automotive manufacturers for making checks to a design to ensure that certain brand- or product-specific targets have been met. The flexibility of Heuristic Analysis means that specific usability criteria, defined by manufacturers for particular products, could be built in to a checklist.

3.5 Layout Analysis

Layout Analysis is a technique used to evaluate an existing interface based on groupings of related functions [24]. It can assist in the restructuring of an interface according to the users' perceived structure of the task. Functions are grouped according to three factors: frequency, importance and sequence of use [24] and a revised design is based on the optimum trade-off between these factors [52]. Layout Analysis was performed for a number of IVIS menu screens, which were identified by the other analytical methods as having usability issues.

3.5.1 IVIS evaluation

A Layout Analysis for one example menu screen is presented to illustrate the process: see figure 7. The MCPA showed that the task time for adjusting fan speed with the remote controller system was reduced when using hard controls, compared with the screen-based controls, which suggested that the design of this menu screen was not optimal. The most significant recommended design change to this menu screen was to reduce the size of the fan speed controls, which had low frequency and importance of use, and to increase the size of the air direction controls, which were used more frequently. In order to make a quantitative comparison between the two input types, the number of layout changes made to each system was used as a metric. The two poorest-performing menu screens for each IVIS (including the remote controller climate screen) were identified according to the results of the other analytic methods. Layout Analysis was performed on the four menu screens and the number of changes recorded. There were eleven changes to the touch screen menus, compared with eighteen for the remote controller menus. This could be an indication that in their current forms, the remote controller menu screens would produce a less effective and efficient interaction than the touch screen menus. However, Layout Analysis is highly subjective and in this study was more useful for producing design recommendations rather than direct comparisons of usability.

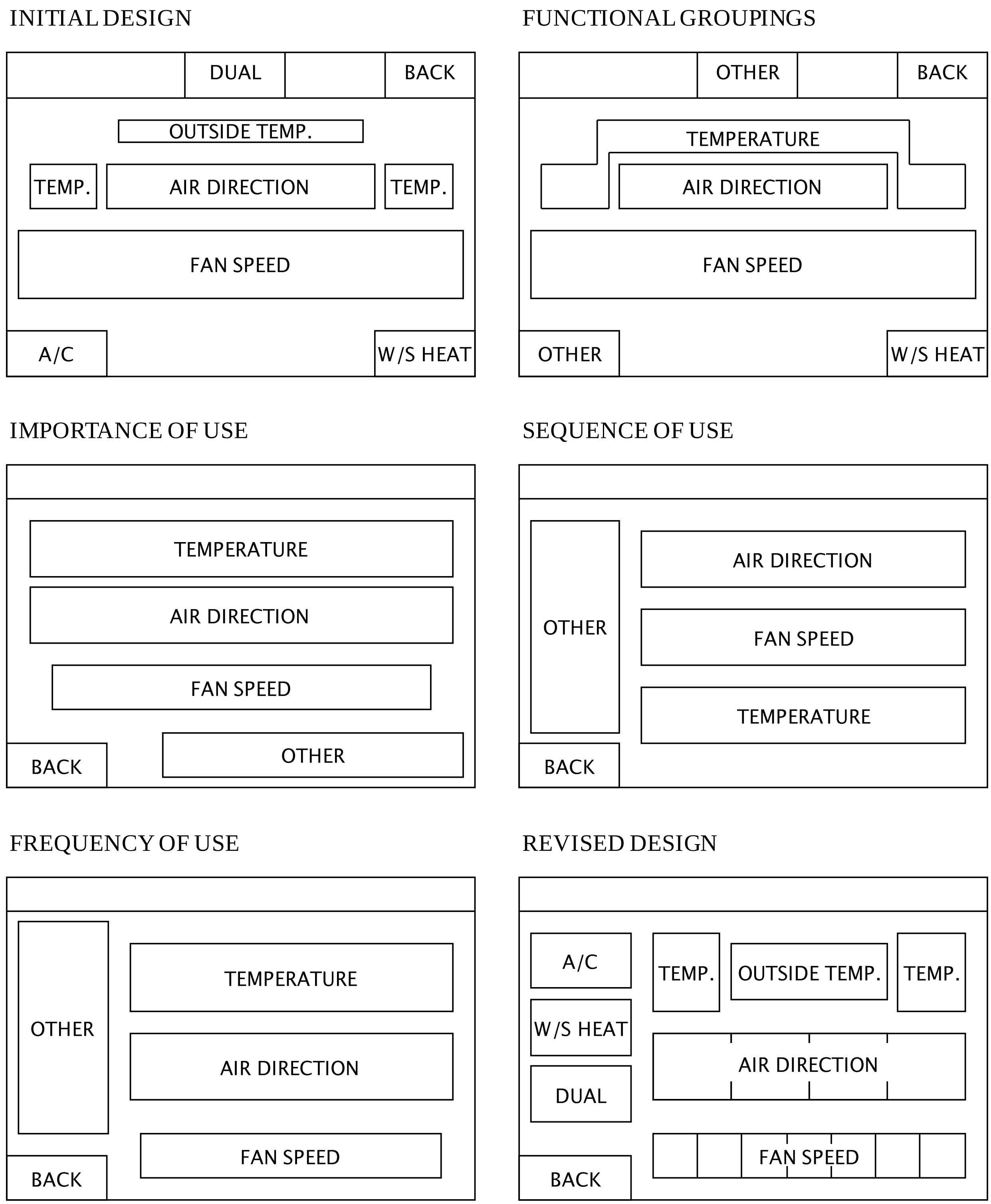

Figure 7. Layout Analysis for the remote controller climate menu.

3.5.2 Utility of the method

Layout Analysis was included in the method set to provide a technique for specifying design changes based on Ergonomics principles. It is not a useful technique for contrasting different systems as number of layout changes is very subjective and is therefore not a valid metric by which to make quantitative comparisons. One use of the technique would be to bridge the gap between evaluation and design: the selection of menus which require redesign is based on the results of the analytic models (MCPA and SHERPA) and the redesign is aimed at addressing the issues identified. Layout Analysis would also be useful at very early stages of design, before the prototyping phase, to assist in initial layout decisions [22].

4. General discussion

The analytic methods applied in this case study were selected to model the performance of two IVIS interfaces. Training, data collection and application times were estimated based on the current study: these will be useful to designers and analysts in future applications of these methods. Training time estimates are presented in table 10 and data collection and analysis times are presented in table 11.

Table 10. Training time estimates for the analytic methods.

Not much time	**Some time**	**Lots of time**
Heuristic Analysis *Familiarisation with checklist* *< 1 hour*	Layout Analysis *Learn layout factors* *1-2 hours*	MCPA *Learn rules, calculation method* *> 2 hours*
		SHERPA *Familiarisation with error codes* *> 2 hours*
		HTA *Learn structure and notation* *> 2 hours*

Table 11. Data collection and application time estimates for the analytic methods.

Not much time	**Some time**	**Lots of time**
Heuristic Analysis *1 hr data collection /* *1 hr analysis*	Layout Analysis *1-2 hrs data collection /* *1 hr analysis per menu screen*	HTA *2-4 hrs data collection /* *6-8 hrs data analysis*
		MCPA *2-4 hrs data collection /* *8-10 hrs data analysis*
		SHERPA *2-4 hrs data collection /* *8-10 hrs data analysis*

These times are similar to those observed in previous studies which have applied these methods [22, 31], with the exception of the application time for Layout Analysis, which was slightly longer than that predicted by Stanton and Young [22]. This is likely to be caused by difference in the interfaces tested in the two studies. Dynamic, screen-based interfaces, analysed in the current study, comprise many different menu layouts in a single system and analysis is therefore likely to be more complex, compared to the static, dashboard mounted controls analysed by Stanton and Young. In comparison with empirical methods which usually require a sample of users interacting with a prototype product, the time and resource requirements of the analytic methods are significantly lower. This supports their application at an early stage in the product development lifecycle. Performance issues identified at this stage can then be further investigated, if necessary, using empirical techniques at a later stage of development when prototypes are more accessible.

Quantitative and qualitative information was extracted from each of the methods in order to make comparisons between the two interaction types under investigation. These data are presented in table 12.

Table 12. Quantitative and qualitative comparisons between the two IVIS.

Method		Touch screen	Remote controller	Best performance?
HTA	Quant.	125 total operations	113 total operations	Remote controller
	Qual.	Most touch screen and remote controller tasks have similar structures but nature of individual operations is different		
MCPA	Quant.	82710 msec total task time	71150 msec total task time	Remote controller
	Qual.	Remote controller times are dependent on the speed of movement through menu options and number of options scrolled through before reaching target. When examined at a task segment level, the touch screen is predicted to produce shorter task times.		
SHERPA	Quant.	6 significant errors	7 significant errors	Touch screen
	Qual.	Remedial measures include increasing the sensitivity and allowing better differentiation between targets for the touch screen; increasing precision of the pointer and moving the enter button for the remote controller.		
Heuristic Analysis	Quant.	7 -ive / 6 +ive issues	8 -ive / 6 +ive issues	Touch screen
	Qual.	Usability issues include glare on the screen and lack of tactile feedback for the touch screen; poor location of back button and complexity of menus for the remote controller.		
Layout Analysis	Quant.	11 layout changes across two menu screens	18 layout changes across two menu screens	Touch screen
	Qual.	In both devices menu targets with highest importance and frequency of use should be placed in the most accessible place on screen. Sequence of use of targets in IVIS interactions should also be accounted for.		

Taking the quantitative data in isolation, it could be concluded that HTA and MCPA support the use of the remote controller over the touch screen and the other three measures, SHERPA, Heuristic Analysis and Layout Analysis, favoured the touch screen over the remote controller. Exploration of the individual methods, however, has shown that it is not sensible to evaluate the IVIS based on this data alone and that some of the methods were unsuitable for making direction comparisons between the touch screen and remote controller interfaces. The findings of this study underline the importance of considering the relevance of outputs on a method-by-method basis [22]: if the results are used solely to identify which system is superior then richer information about wider aspects of usability could be lost. Gray and Salzman [19] warned that the advantages and disadvantages of analytic methods must be understood in order to mitigate against erroneous claims about system usability.

5.3.6.1. Analytic Methods for IVIS evaluation

HTA produced a hierarchical outline of tasks, which described the smallest operations which a user performs when interacting with a particular interface. This analysis showed that the basic task segments for selecting a menu target consisted of the same number of operations for both systems and highlighted the effect of task structure on interaction strategies with the two IVIS, i.e. the touch screen generally required more target presses to complete tasks than the remote controller. Operations identified by the HTA were then fed in to the MCPA and assigned duration times in order to calculate predictions of total task times. Like the HTA, MCPA also highlighted the differences in task structure between the two IVIS; however, the MCPA also showed that although the number of operations in a task segment was consistent, the operation timings assigned to these operations produced differences between the task times of the touch screen and remote controller. Although there was some overlap between the output of these two methods, both are recommended in IVIS evaluation. HTA is a necessary precursor for other methods, including MCPA and SHERPA, and is thought to be a useful exercise for familiarising designers and evaluators with task structures. MCPA expands the output of HTA by assigning predicted times to the tasks and the task time metric is useful for comparing IVIS, and for making estimates about the effect of IVIS tasks on concurrent tasks, such as driving. SHERPA highlighted a number of potential errors with both systems which would be useful to a designer at the early stages of product development; the remedial strategies devised as part of the analysis would guide any necessary redesign activities in order to reduce errors in the driver-IVIS interaction. There was, however, quite significant overlap between the errors which were identified for the two systems, which does not support the use of SHERPA as a comparative evaluation tool. SHERPA is based on an objective task description and the analysis follows a rigid structure that produces quantifiable results; however the assignment of error frequency and severity is dependent on the analyst’s judgement. The remedial strategies recommended as part of SHERPA are also an example of qualitative output. Comparison of heuristic analysis with the results of SHERPA, in a process of data triangulation [19, 36, 53, 54], showed that both methods identified some of the same usability issues; however, SHERPA errors tended to relate to individual operations and issues which may prevent these being performed successfully, whereas the heuristic analysis identified more general issues relating to the system and wider environment, e.g. glare and reflections obscuring the display screen. There were also instances where the two methods did not agree and this has also been found in previous studies [e.g. 55]. For example, glare on the touch screen would lead to a R1 SHERPA error (information not obtained); however, the SHERPA method, which is based on the HTA specification, does not support the analyst in accounting for environmental factors and therefore this was not identified as a potential error. The issue of false positive error detection has also been found in studies of SHERPA [31, 55]: this could encourage unnecessary changes to a design. Heuristic analysis identifies usability issues and the assumption is that these will lead to poor usability when the IVIS is used by consumers; however, identification of usability issues is not a guarantee of poor performance [19]. This problem is compounded by the lack of information about frequency of occurrence of issues in this type of analysis. Layout analysis was only applied to the two worst-performing menu screens in both IVIS; therefore it is very difficult to make quantitative comparisons between the touch screen and remote controller based on this information alone. The subjectivity of techniques like Layout Analysis, and also Heuristic Analysis and SHERPA, is a disadvantage in situations where quantifiable metrics are needed so that two or more competing systems can be compared. These techniques also suffer from problems associated with the assumption that the analyst always has implicit knowledge of the context-of-use [56]: this is often not the case. However, Layout Analysis still adds to the analytic approach by providing a strategy for exploring existing GUI layouts: this is important as the GUI should be optimised with task structure and input device to produce ideal system performance. It also provides designers with a structured method for addressing the types of usability issues identified by SHERPA and Heuristic Analysis.

5.3.6.2. Context Versus Objectivity

Usability evaluation should account for the specific context within which systems are used [18]; however, the results showed that not all of the methods addressed this issue. HTA and MCPA were developed for application in a single task environment, which means that in this case the effects of driving on IVIS effectiveness were not modelled. Based on this case study, it appeared that the more a method accounts for the broad effects of context, the more subjective it becomes. On the other hand, a narrow and more objective focus produces quantitative models, which enable direct comparisons between systems to be made [56]. For example, MCPA allows detailed, quantitative, comparable predictions for a very specific aspect of usability; however, the focus on only one aspect of system effectiveness (task times in a single task environment), means that contextual factors are not accounted for [57]. Subjective techniques enable a broader approach, which aims to capture the 'whole essence of user-device interaction' [22], and these methods therefore account for context to some extent. However, the qualitative nature of the outputs means that these methods do not drill down to a deep level of detail and are therefore more suited to usability checks (e.g. heuristic analysis) or design recommendations (e.g. layout analysis and SHERPA), rather than direct comparisons [38, 58-63].

5.3.6.3. Extending MCPA

To address the trade-off between context and objectivity an extension to MCPA which allows consideration of the context-of-use is proposed. MCPA measures performance via quantitative predictions of task time rather than relying on the assumption that poor performance will follow on from identification of usability issues [19, 51, 57]. Another advantage of MCPA is that it takes a taskonomic approach to modelling HMI [64], which means that systems are analysed in terms of the activity or task being performed. On the other hand, the heuristic checklist applied in this study took a taxonomic approach because it analysed elements of an interface based on functional, rather than task-based, categories [22, 36]. Nielsen [64] argued that both taxonomies and taskonomies are necessary in design; however, in a dual-task driving context, where interaction with secondary tasks is so dependent on the concurrent demand from driving, the activity-based approach [36] appears to be the most useful for usability evaluation. MCPA in particular has potential for analysing these dual-task interactions because the driver's interaction with primary driving tasks can be incorporated into the models in parallel to IVIS operations. This technique could be used as a direct measure of the effectiveness of the user-system interaction in a dual task driving environment.

5. Conclusions

The aim of the case study presented in this paper was to explore an analytic approach to IVIS usability modelling to meet a requirement for early-stage, low resource product evaluation. The methods were selected to model important aspects of HMI performance: task structure, interaction times, error rates, usability issues, and interface design [19]. The findings of the study have been discussed in terms of IVIS comparisons, utility of the methods, time and resource demands, and potential for further development. HTA was not useful for making relative comparisons between systems; however, it was found to be an essential starting point for MCPA and SHERPA and was also useful for the exploration of task structure. MCPA modelled task interaction times as a measure of performance; however, in its current state it does not account for the dual task driving scenario. There is however, potential to extend the method to address this issue. SHERPA was expected to yield a comprehensive list of potential errors guided by its structured taxonomic approach; however, assessment of error frequency and severity are still largely open to analyst bias. Data triangulation against the results of the heuristic analysis also showed that neither method was comprehensive. Heuristic Analysis is not suitable for comparisons between systems; however, there is potential for development as product- or brand-specific guidance. Heuristic Analysis also has an advantages of low training and application times, which supports its use for early identification of potential usability issues. Layout Analysis appears to be useful for bridging the gap between evaluation and design and has only moderate time and resource demands, which will enable analysts to not only make quick decisions about product performance but also to make recommendations to improve usability. The findings of this exploratory study have highlighted a trade-off between subjectivity and focus on context-of-use. An extension of the MCPA modelling method has been suggested to incorporate analysis of context into a quantitative technique so that more useful predictions of IVIS performance can be made. This will be a focus of future work.

References

[1] Fastrez, P. and Haué, J.-B.: 'Editorial: designing and evaluating driver support systems with the user in mind', International Journal of Human-Computer Studies, 2008, 66, (3), pp. 125-131

[2] Hedlund, J., Simpson, H. and Mayhew, D.: 'International conference on distracted driving: summary of proceedings and recommendations', (The Traffic Injury Research Foundation and The Canadian Automobile Association, 2006)

[3] Young, K.L., Regan, M.A. and Lee, J.D.: 'Factors moderating the impact of distraction on driving performance and safety', in Regan, M.A., Lee, J.D. and Young, K.L. (ed.): 'Driver distraction: theory, effects and mitigation' (CRC Press, 2009, 1st edn.), pp. 335-351

[4] Gu Ji, Y. and Jin, B.S.: 'Development of the conceptual prototype for haptic interface on the telematics system', International Journal of Human-Computer Interaction, 2010, 16, (1), pp. 22-52

[5] Krems, J.F.: 'Editorial: selected papers from the 2nd European conference on human centred design for intelligent transport systems', IET Intelligent Transport Systems, 2011, 5, (2), pp. 101-102

[6] Mitsopoulos-Rubens, E., Trotter, M.J. and Lenné, M.G.: 'Usability evaluation as part of iterative design of an in-vehicle information system', IET Intelligent Transport Systems, 2011, 5, (2), pp. 112-119

[7] Pickering, C.A., Burnham, K.J. and Richardson, M.J.: 'A review of automotive human machine interface technologies and techniques to reduce driver distraction'. Proc. 2nd IET International Conference on System Safety, London, 22-24 October, 2007, pp. 223-228

[8] Pereira, M., Hamama, H., Dapzol, N., Bruyas, M.P. and Simões, A.: 'Simultaneous interaction with in-vehicle systems while turning left: comparison among three groups of drivers', IET Intelligent Transport Systems, 2009, 3, (4), pp. 369-378

[9] Miller, R.B.: 'Human ease of use criteria and their tradeoffs', (IBM Corporation, 1971)

[10] Dehnig, W., Essig, H. and Maass, S.: 'The adaptation of virtual man-computer interfaces to user requirements in dialogs' (Springer-Verlag, 1981, 1st edn.)
[11] Dzida, W., Herda, S. and Itzfeldt, W.D.: 'User-perceived quality of interactive systems'. Proc. 3rd International Conference on Software Engineering, Atlanta, GA, 10-12 May, pp. 188-195
[12] Shackel, B.: 'Ergonomics in design for usability'. Proc. 2nd Conference of The British Computer Society Human Computer Special Interest Group, York, 23-26 September, 1986, pp. 44-64
[13] Norman, D.A.: 'Design principles for human computer interfaces'. Proc. SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, 12-15 December, 1983, pp. 1-10
[14] Nielsen, J.: 'Usability engineering' (Academic Press, 1993, 1st edn.)
[15] Shneiderman, B.: 'Designing the user interface: strategies for effective human-computer interaction' (Addison-Wesley, 1992, 2nd edn.)
[16] Bevan, N.: 'International standards for HCI and usability', International Journal of Human-Computer Studies, 2001, 55, (4), pp. 532-552
[17] ISO 9241: 'Ergonomic requirements for office work with visual display terminals (VDTs) - part 11: guidance on usability', 1998
[18] Harvey, C., Stanton, N.A., Pickering, C.A., McDonald, M. and Zheng, P.: 'A usability evaluation toolkit for in-vehicle information systems (IVISs)', Appl. Ergon, 2011, 42, (4), pp. 563-574
[19] Gray, W.D. and Salzman, M.C.: 'Damaged merchandise? A review of experiments that compare usability evaluation methods', Human-Computer Interaction, 1998, 13, (3), pp. 203-261
[20] Greenberg, S. and Buxton, B.: 'Usability evaluation considered harmful (some of the time)'. Proc. CHI 2008 Conference on Human Factors in Computing Systems, Florence, 5-10 April, 2008, pp.
[21] Broström, R., Bengtsson, P. and Axelsson, J.: 'Correlation between safety assessments in the driver-car interaction design process', Applied Ergonomics, 2011, 42, (4), pp. 575-582
[22] Stanton, N.A. and Young, M.S.: 'A guide to methodology in ergonomics: designing for human use' (Taylor and Francis, 1999, 1st edn.)
[23] Kirwan, B. and Ainsworth, L.K.: 'A guide to task analysis' (Taylor and Francis, 1992, 1st edn.)
[24] Stanton, N.A., Salmon, P.M., Walker, G.H., Baber, C. and Jenkins, D.P.: 'Human factors methods: a practical guide for engineering and design' (Ashgate, 2005, first edn.)
[25] Stanton, N.A.: 'Hierarchical task analysis: developments, applications, and extensions', Applied Ergonomics, 2006, 37, (1), pp. 55-79
[26] Baber, C. and Mellor, B.: 'Using critical path analysis to model multimodal human-computer interaction', Int. J. Hum. Comput. St, 2001, 54, (4), pp. 613-636
[27] Wickens, C.D.: 'Processing resources and attention ', in Damos, D.L. (ed.): 'Multiple-task performance' (Taylor and Francis, 1991, 1st edn.), pp. 3-34

[28] Harrison, A.: 'A survival guide to critical path analysis' (Butterworth-Heinemann, 1997, 1st edn.)

[29] Green, P.: 'The 15-second rule for driver information systems'. Proc. ITS America Ninth Annual Meeting, Washington, DC, 9th April, 1999, pp. 1-9

[30] Nowakowski, C., Utsui, Y. and Green, P.: 'Navigation system destination entry: the effects of driver workload and input devices, and implications for SAE recommended practice', (University of Michigan Transportation Research Institute (UMTRI), 2000)

[31] Baber, C. and Stanton, N.A.: 'Human error identification techniques applied to public technology: predictions compared with observed use', Applied Ergonomics, 1996, 27, (2), pp. 119-131

[32] Lyons, M.: 'Towards a framework to select techniques for error prediction: supporting novice users in the healthcare sector', Applied Ergonomics, 2009, 40, (3), pp. 379-395

[33] Nowakowski, C. and Green, P.: 'Prediction of menu selection times parked and while driving using the SAE J2365 method', (University of Michigan Transportation Research Institute (UMTRI), 2001)

[34] SAE J2365 (proposed draft): 'Calculation of the time to complete in-vehicle navigation and route guidance tasks', 2002

[35] Stevens, A., Board, A., Allen, P. and Quimby, A.: 'A safety checklist for the assessment of in-vehicle information systems: a user's manual', (Transport Research Laboratory, 1999)

[36] Wilson, C.E.: 'Triangulation: the explicit use of multiple methods, measures, and approaches for determining core issues in product development', Interactions, 2006, pp. 46-63

[37] Stanton, N.A. and Baber, C.: 'Modelling of human alarm handling response times: a case study of the Ladbroke Grove rail accident in the UK', Ergonomics, 2008, 51, (4), pp. 423-440

[38] Olson, J.S. and Moran, T.P.: 'Mapping the method muddle: guidance in using methods for user interface design', in Rudisill, M., Lewis, C., Polson, P.G. and MacKay, T.D. (ed.): 'Human-computer interface designs: success stories, emerging methods, and real world context' (Kaufmann, 1996, edn.), pp. 269-300

[39] Wickens, C.D.: 'Multiple resources and performance prediction', Theor. Issues Ergonomics Sci, 2002, 3, (2), pp. 159-177

[40] Baber, C.: 'Critical path analysis for multimodal activity', in Stanton, N., Hedge, A., Brookhuis, K., Salas, E. and Hendrick, H. (ed.): 'Handbook of human factors and ergonomics methods' (CRC Press, 2005, 1st edn.), pp. 41-41 - 41-48

[41] Olson, J.R. and Olson, G.M.: 'The growth of cognitive modeling in human-computer interaction since GOMS', Human-Computer Interaction, 1990, 5, (2), pp. 221-265

[42] Card, S.K., Moran, T.P. and Newell, A.: 'The psychology of human-computer interaction' (Lawrence Erlbaum, 1983, first edn.)

[43] Ackerman, P.L. and Cianciolo, A.T.: 'Psychomotor abilities via touch-panel testing: measurement innovations, construct, and criterion validity', Human Performance, 1999, 12, (3-4), pp. 231-273
[44] Rogers, W.A., Fisk, A.D., McLaughlin, A.C. and Pak, R.: 'Touch a screen or turn a knob: choosing the best device for the job', Human Factors, 2005, 47, (2), pp. 271-288
[45] Card, S.K., English, W.K. and Burr, B.J.: 'Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT', Ergonomics, 1978, 21, (8), pp. 601-613
[46] Alliance of Automobile Manufacturers. 'Statement of principles, criteria and verification procedures on driver interactions with advanced in-vehicle information and communication systems', (AAM, 2006)
[47] Bhise, V.D., Dowd, J.D. and Smid, E.: 'A comprehensive HMI evaluation process for automotive cockpit design'. Proc. SAE World Congress, Detroit, MI, March 3-6, 2003, pp. 61-70
[48] Commission of the European Communities. 'Commission reccomendation of 26/V/2008 on safe and efficient in-vehicle information and communication systems: update of the European statement of principles on human-machine interface', (Commission of the European Communities, 2008)
[49] Japan Automobile Manufacturers Association. 'Guideline for in-vehicle display systems, version 3.0', (JAMA, 2004)
[50] Nielsen, J. Ten usability heuristics. 2005 [cited 2009 20.03.09]; Available from: http://www.useit.com/papers/heuristic/heuristic_list.html.
[51] Mendoza, P.A., Angelelli, A. and Lindgren, A.: 'Ecological interface design inspired human machine interface for advanced driver assistance systems', IET Intelligent Transport Systems, 2011, 5, (1), pp. 53-39
[52] Stanton, N.A. and Young, M.S.: 'Is utility in the mind of the beholder? A study of ergonomics methods', Appl. Ergon., 1998, 29, (1), pp. 41-54
[53] Hancock, P.A. and Szalma, J.L.: 'On the relevance of qualitative methods for ergonomics', Theoretical Issues in Ergonomics Science, 2004, 5, (6), pp. 499-506
[54] Mackay, W.E.: 'The interactive thread: exploring methods for multi-disciplinary design'. Proc. 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, Cambridge, MA, 1-4 August, 2004, pp. 103-112
[55] Stanton, N.A. and Stevenage, S.V.: 'Learning to predict human error: issues of acceptability, reliability and validity', Ergonomics, 1998, 41, (11), pp. 1737-1756
[56] Blandford, A. and Rugg, G.: 'A case study on integrating contextual information with analyltical usability evaluation', International Journal of Human-Computer Studies, 2002, 57, (1), pp. 75-99
[57] Bevan, N. and Macleod, M.: 'Usability measurement in context', Behaviour and Information Technology, 1994, 13, (1), pp. 1994

[58] Burns, P.C., Trbovich, P.L., Harbluk, J.L. and McCurdie, T.: 'Evaluating one screen / one control multifunction devices in vehicles'. Proc. 19th International Technical Conference on the Enhanced Safety of Vehicles, Washington, DC, 6-9 June, 2005, pp. 1-9

[59] Butters, L.M. and Dixon, R.T.: 'Ergonomics in consumer product evaluation: an evolving process', Applied Ergonomics, 1998, 29, (1), pp. 55-58

[60] Cherri, C., Nodari, E. and Toffetti, A.: 'Review of existing tools and methods: AIDE deliverable 2.1.1.', (Volvo Technology Corporation (VTEC), 2004)

[61] Jeffries, R., Miller, J.R., Wharton, C. and Uyeda, K.M.: 'User interface evaluation in the real world: a comparison of four techniques'. Proc. CHI 1991 Conference on Human Factors in Computing Systems, Pittsburgh, PA, 15-20 May, 1999, pp. 119-124

[62] Nielsen, J.: 'Finding usability problems through heuristic evaluation'. Proc. CHI 1992 Conference on Human Factors in Computing Systems, Monterey, CA, 3-7 May, 1992, pp. 373-380

[63] Nielsen, J. and Phillips, V.L.: 'Estimating the relative usability of two interfaces: heuristic, formal, and empirical methods compared'. Proc. CHI 1993 Conference on Human Factors in Computing Systems, Amsterdam, 24-29 April, 1993, pp. 214-221

[64] Nielsen, J.: 'Logic versus usage: the case for activity-centered design', Interactions, 2006, pp. 45