EXPLORING MEASURES OF USABILITY FOR IN-VEHICLE TECHNOLOGY

Alan Stevens - Chief Scientist, Transportation, TRL
Alistair Weare - Human Factors Researcher, TRL

ABSTRACT:

This paper examines various tools and information resources available to designers and usability professionals when developing or evaluating the HMI (Human Machine Interaction) of IVIS (In-Vehicle Information Systems) with regard to usability. It starts with a consideration of what the word ‘usability’ means and how it has been defined in the literature, along with a discussion of how it is relevant to IVIS. The paper then reviews HMI guidelines that have been produced and are in current usage, and finally looks at some of the usability assessment methods available to HMI professionals.

1 INTRODUCTION (WHAT IS USABILITY?)

The late 1970’s and early 80’s saw the arrival of personal computers in the public domain, and kick-started the movement towards ‘usability’ as a design consideration. Whilst usability is now an accepted concept, it is not rigidly defined (though definitions have certainly been proposed). This section will review what is meant by usability, and which aspects are of greatest importance when considered in the context of IVIS, specifically with regard to the HMI of IVIS.

Nielsen [1], whilst not necessarily defining usability, proposed five key usability attributes for products:

- Learnability
- Efficiency
- Memorability
- Errors
- Satisfaction

One of the most widely quoted definitions of usability is from ISO 9241 [2], which takes three of the attributes from Nielsen to define 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'

The first two of the three concepts can be broadly defined as follows:

- Effectiveness – extent to which a product does what it was designed to do
- Efficiency – resources required to achieve a task

However, satisfaction as a concept is much less tangible and, indeed, is a highly subjective construct. The ISO definition also makes no allowance for the dimension of time or the idea that prolonged exposure to a product or system may change a user's perception of usability. Jordan [3] addressed this with his proposal of five higher-order components within the concept of usability:

- Guessability
- Learnability
- Experienced User Potential (EUP)
- System potential
- Re-usability (or memorability)

Jordan's concept places a much greater emphasis on how the user experiences the product/system over the course of its use; from initial access through to the experienced user.

This raises the question of whether usability is an inherent property of the product/system or an outcome of its use; a question raised more explicitly by Bevan [4]. This, perhaps, is of particular interest in relation to in-vehicle systems, because use of a particular system in isolation is not necessarily of great importance. The focus for in-vehicle technologies is, instead, typically on how that technology will affect the safety of the overall driving task. ISO 17287 [5] defined a new and related concept - 'suitability' - as:

'The degree to which a [system] is appropriate in the context of the driving environment based on compatibility with the primary driving task'

The usability of a particular product has therefore to consider carefully the context of its use. A user may well regard a product as being extremely usable, but fail to consider fully how use of that product may interact with the driving task. If a product is potentially to be used within a vehicle the designer must first understand the user and the context of use (driving experience, technology experience, expectations and so on).

Incorporating the earlier concepts of satisfaction and acceptability leads to a consideration of both the usability and the utility of a product/system in order to understand how it will be received (and used) by a user. Indeed, the Technology Acceptance Model (TAM) [6] describes how perceived usefulness and ease of use are the main determinants of attitude towards a technology, which in turn predicts intention to use and, ultimately, actual system use. It can therefore be seen as the responsibility of the designer to ensure that systems are not only perceived as easy to use but are, in fact, usable within the driving context.

In determining the usability of an IVIS (whether designed specifically for use in a vehicle or not), there must be an understanding of how the system fits into the larger vehicle-driver-road system. It must be useful to the driver within the higher driving task, efficient such that it presents a minimal distraction, and its ease of use must be compatible with any competing demands on the driver at the time of use (which may or may not be when driving).

2 DESIGN GUIDELINES

This section reviews a range of standards and guidelines, available to designers, that aim to promote usability in the driving context.

2.1 Regulations and Standards

2.1.1 International regulations

International standards may not be legally binding, but do form a framework within which designers can seek to create products and systems that subscribe to a common philosophy. Standards, by their nature of attempting to define best practice, are often referenced in national regulations or supply contracts and so often influence mandatory requirements. As such they have an important role to play, but only if they are kept up to date. At least three

ISO groups are currently working in areas relevant to IVIS and usability:

- ISO TC 22 SC13 WG8 covers basic standards for human factors design of in-vehicle systems;
- ISO TC 204 WG14 concerns vehicle and cooperative services (and some interface issues) including, for example, Lane Departure Warning and automatic Emergency Braking Systems; and
- ISO TC 204 WG17 concerns nomadic and portable devices for ITS services.

A multitude of standards have been produced that cover the design of visual and audible driver interfaces, much of which has formed the basis for current design guidelines and codes of practice.

2.1.2 United States regulations

In the US, laws about in-vehicle distraction generally fall under the jurisdiction of individual states but with some at the national (federal) level. As an example of state provision, the US state of Nevada passed a law in June 2011 concerning the operation of driverless (fully automated) cars whereby the Nevada Department of Motor Vehicles is responsible for setting safety and performance standards and for designating areas where driverless cars may be tested.

As an example of national provision, in October 2009 President Obama issued an Executive Order prohibiting Federal employees from texting while driving. This order is specific to employees' use of Government owned vehicles, or privately owned vehicles while on official Government business, and includes texting-while-driving, and using wireless electronic devices supplied by the Government.

2.1.3 European regulations

There is currently little in the way of European legislation specifically related to the HMI of IVIS. However, the European Commission published a Directive in 2010 [7], which has provisions for the development of specifications and standards for ITS road safety including HMI and the use of nomadic devices. European regulations may be a consideration in the future.

2.2 Design Guidelines

2.2.1 Europe: European Statement of Principles

The European Commission (EC) [8] has supported the development of a document called the ‘European Statement of Principles on HMI’ (referred to as ESoP) which provides high-level HMI design advice. As an EC Recommendation it has the status of a recommended practice or Code of Practice for use in Europe. The EC Recommendation also contains 16 Recommendations for Safe Use (RSU), which build on Health and Safety legislation by emphasising the responsibility of organisations that employ drivers to attend to HMI aspects of their workplace. Adherence to the RSU is intended to promote greater acceptance of technology by drivers.

The design-guidelines part of the ESoP comprises 34 principles to ensure safe operation while driving. These are grouped into the following areas: Overall Design Principles, Installation Principles, Information Principles, Interactions with Controls and Displays Principles, System Behaviour Principles and Information about the System Principles.

2.2.2 United States: Alliance and NHTSA

The US motor vehicle manufacturers have developed ‘Alliance Guidelines’ that cover similar, high-level, design principles to the ESoP. The Guidelines [9] consist of 24 principles organised into five groups: Installation Principles, Information Presentation Principles, Principles on Interactions with Displays/Controls, System Behaviour Principles, and Principles on Information about the System.

The USA’s National Highway Transportation Safety Administration (NHTSA) has worked with the auto industry and the cell phone industry to develop a set of guidelines [10] for visual-manual interfaces for in-vehicle technologies. These are based on the ESoP/Alliance guidelines and introduce some specific assessment procedures. The NHTSA plan to publish guidelines for portable devices and for voice interfaces in future years.

The NHTSA guidelines seek to provide specific acceptance criteria for the given design principles, as opposed to the more generic criteria given within the ESoP. Whilst this does lead to more definitive assessment, it does rely on the testing of participants to determine levels of distraction (suggesting the use of 24 people to test). This introduces issues of participant homogeneity and sample sizes and, as NHTSA notes, this may mean that outcomes may differ between different test groups.

2.2.3 Japan: JAMA

The Japanese Auto Manufacturers Association (JAMA) Guidelines [11] consist of four basic principles and 25 specific requirements that apply to the driver interface of each device to ensure safe operation while driving. Specific requirements are grouped into the following areas: Installation of Display Systems, Functions of Display Systems, Display System Operation While Vehicle in Motion, and Presentation of Information to Users. Additionally, there are three annexes: Display Monitor Location, Content and Display of Visual Information While Vehicle in Motion, and Operation of Display Monitors While Vehicle in Motion. There is, as well, one appendix: Explanation of the guideline for in-vehicle display systems.

2.3 Warning Guidelines

Guidelines on establishing requirements for high-priority warning signals have been under development for more than five years by the UNECE/WP29/ITS Informal Group [12]. There has also been work in standardisation groups to identify how to prioritise warnings when multiple messages need to be presented and one ‘Technical specification’ (TS) has been produced:

- ISO/TS 16951: Road Vehicles – Ergonomic aspects of transport information and control systems – Procedures for determining priority of on-board messages presented to drivers

In addition, two Technical Reports are relevant that contain a mixture of general guidance information, where supported by technical consensus, and discussion of areas for further research:

- ISO/PDTR 16352: Road Vehicles – Ergonomic aspects of transport information and control systems – MMI of warning systems in vehicles
- ISO/PDTR 12204: Road Vehicles – Ergonomic aspects of transport information and control systems – Introduction to integrating safety critical and time critical warning signals

2.4 Driver Assistance Systems Guidelines

To help promote acceptance of Advanced Driver Assistance Systems (ADAS), a key issue is ensuring controllability and this has been addressed through guidelines. Controllability is determined by the possibility and driver's capability to perceive the criticality of a situation; the driver's capability to decide on appropriate countermeasures (such as overriding or switching off the system) and the driver's ability to perform any chosen countermeasures (taking account of the driver's reaction time, sensory-motor speed and accuracy). Drivers will expect controllability to exist in all their interactions with assistance systems:

- during normal use within system limits
- at and beyond system limits
- during and after system failures

The European project RESPONSE developed a Code of Practice for defining, designing and validating ADAS. The Code describes current procedures used by the vehicle industry to develop safe ADAS with particular emphasis on the human factors requirements for 'controllability'.

Another European project, ADVISORS, attempted to integrate the RESPONSE Code within a wider framework of user-centred design taking account of the usability of information, warning and assistance systems [13]. There is also activity by the International Harmonized Research Activities – Intelligent Transport Systems (IHRA-ITS) Working Group to develop a set of high-level principles for the design of driver assistance systems [14].

3 METHODS RELATED TO USABILITY MEASUREMENT

As discussed earlier, when assessing usability of IVIS it is not necessarily the inherent usability of the product/system in isolation that is of interest; it is usually more important to gauge the usability of the system within the wider context of the driving task as a whole. As such, metrics of usability need to focus on overall driver performance associated with system use. Also, as shown in the previous section, various guidelines have been produced that provide designers and assessors with principles that are shown to help promote better system performance. Methods for assessing IVIS usability can therefore take two different approaches:

1. Evaluate actual driver performance when using the product/system in a realistic context of use
2. Evaluate how well a product/system meets the design principles in the relevant guidelines

With regard to approach 1 (measuring actual performance) a method for testing the usability of in-car systems can be seen to be a combination of three factors [15]:

1. Which environment the method is used in (road, test track, simulator, laboratory etc.) The decision will be based partly on the available resources; on-road testing is usually far more expensive than laboratory testing, for example. More importantly, from a scientific point of view, is to do with the validity of the testing, and there is usually a trade-off between achieving good ecological or internal validity. For instance, road trials may have high ecological validity (we are confident that the testing is representative of the real world), but may have poor internal validity (it may not be clear what the exact reasons are for any observed behaviours). Laboratory testing, conversely, will typically demonstrate good cause-effect relationships, but it is not always clear if such relationships are relevant or occur in practice in the real world.
2. Which task manipulations occur (multiple task, single task loading, no tasks given etc.) When testing IVIS, there is usually a need to, at least partially, recreate relevant operational tasks, in order to make usual judgements on usability and performance. The broadness of this representation will depend on the research/evaluation needs.
3. Which dependent variables (operationalised as metrics) are of interest. Choosing very specific metrics (such as eyes-off-road time) can give very objective performance measures that allow clear comparison between conditions. Conversely, choosing much broader metrics (such as questionnaire self-assessment of task difficulty) are

typically much more subjective and less focussed, but may give a better overall picture of performance, typically a combination of measures will be used and will depend on the research questions to be answered.

As noted by Rogers et al. [16], in deciding on any method, the design team must consider the overall goals of the work, specific questions to be addressed, the practical and ethical issues, and how data will need to be analysed and reported. Due to the unique combinations of complex tasks demands when driving, a number of bespoke research methods have arisen that are typically used in assessments within a driving context. In preparation for the NHTSA guidelines, Ranney et al. [17] assessed several different methods for testing distraction potential of devices. The following are some commonly used methods for testing usability:

- Road trials – Participants drive a real vehicle, usually on the public roads but potentially on a test track. Typically a wide range of more generic metrics are used in order to understand what will usually be a complex set of observed behaviours, causal/contributory factors and driving environments.
- Simulator trials – participants drive in a simulated environment in which the environment is carefully controlled and where test activities are tightly controlled. The key for simulator trials is the repeatability of the experiments. Specific objective measures allow useful comparisons between participants and between test scenarios.
- Occlusion - This is a standardised laboratory-based method [18] which focuses on the visual demand of in-vehicle systems. It involves the use of special goggles that mimic periodic glancing behaviour typically adopted by drivers when attempting to perform a secondary visual task. Common performance metrics are: time taken to complete task, number of glances (related) and number of errors made.

- Peripheral detection – This method is often performed as part of a simulator study and requires participants to respond to changes in their periphery. This may be the presence of lights or movement of shapes. Speed and accuracy of responses are used as metrics and as an indication of mental workload and distraction associated with secondary tasks.

- Lane change task - This standardised method [19] uses a basic PC simulated environment in which drivers are requested to make various lane change manoeuvres whilst engaging with an in-vehicle system. The extent to which the profile of manoeuvre made by a driver varies from the optimum manoeuvre (the normative model) is considered to be a measure of the quality of their driving.

- Keystroke Level Model (KLM) - The KLM method is a form of task analysis in which system tasks with a given user-interface are broken down into their underlying physical and mental operations; e.g., pressing buttons, moving hand between controls, scanning for information. Time values are associated with each operator and summed to give a prediction of task times. In an extension of the KLM method, Pettitt, Burnett and Stevens [20] have developed new rules that enable designers to develop predictions for a range of visual demand measures.

The second main option for evaluating IVIS is to develop checklists based on the design principles laid down in the guidelines (e.g. ESoP, NHTSA and JAMA guidelines). To this end a functional IVIS usability checklist has been developed [21], based on an existing checklist produced for the UK Government in the late 1990's, and incorporates requirements taken from the ESoP.

The structure of the checklist is such that each specific topic comprises an initial question, supplemented by optional sub-questions requiring True/False answers and then a response box identifying if there are concerns/issues about the interface design in relation to the question (Figure 1).

Is the driver's view of the road scene free from obstruction by the IVIS?

The swept windscreen area is fully clear.	True/False/NA
The view of the mirroes is not restricted.	True/False/NA
The side windows are fully clear.	True/False/NA

None		Minor		Serious		N/A	

Fig. 1 Example question from the TRL Checklist

A key issue in developing the Checklist was the extent to which questions are sufficiently "elemental" such that they can be answered by observation and without judgment (i.e. such that all reasonable observers would agree on the same answer). For just three Checklist questions derived from the ESoP principles, it was proposed that further work on measurements and criteria needs to be established:

- Is the IVIS securely fitted?
- Is the IVIS visual display positioned close to the driver's normal line of sight?
- Are presented messages visually simple?

A future challenge is to consider whether the checklist can be developed to include a rating system, whereby differing systems can be compared in terms of their overall usability. There are, however, a number of difficult issues that would need to be addressed:

- Elements - Which elements are included within the rating? (E.g. all Checklist questions)
- Scoring* - How are the individual elements scored? (E.g. +3/0/-3 or 1-10)
- Weighting - How are the individual elements weighted? (E.g. all

even, high and low weights, individual weights)

- Combining - How are the scores and weights combined?
- Rating - How is the final number converted into the consumer rating?

*An additional issue related to scoring is determining how to account for features that may be absent. Is it “better” to have a feature, even if poorly designed, than for that feature to be absent; and how should the scoring reflect this?

4 CONCLUSIONS

This paper has discussed how usability can usefully be considered in terms of perceived usefulness and ease of use, which contribute to a drivers' judgement of the acceptability of in-vehicle technology and their desire to use it. It has considered the importance of context in terms of assessing usability, particularly from a safety perspective.

The paper has reviewed a range of regulations, standards and design guidelines that aim to encourage better-designed in-vehicle technology that should also help to promote driver acceptance. Although basic human factors principles are established, the rapid development of in-vehicle technology presents a challenge for updating regulations and detailed design guidance.

Finally, the paper has explored a range of methods through which usability can be evaluated. The technique, the equipment used and the testing environment need to be carefully chosen depending on the in-vehicle system and the evaluation question being addressed. Some questions remain in terms of producing usability ratings suitable for the public domain; nevertheless, it can be concluded that usability can be measured and that usability is a key contributor to drivers' acceptance of in-vehicle technology.

5 REFERENCES

1. Nielsen, J.: 'Usability Engineering' (Margan Kaufmann: San Diego; 1993)
2. ISO/IEC 9242-11: 'Ergonomic requirements for office work with visual display terminals (VDT)s – Part 11: Guidance on usability', 1998
3. Jordan, P.W.: 'Introduction to usability' (Taylor and Francis: London, 1998)

4. Bevan, N.: 'International standards for HCI and usability', International Journal of Human-Computer Systems, 2001, 55(4), pp. 533-552

5. ISO 17287: 'Road vehicles – Ergonomic aspects of transport information and control systems – Procedure for assessing suitability for use while driving', 2003

6. Davis, F.D.,Bagozzi, R.P. and Warshaw, P.R.: 'User acceptance of computer technology: a comparison of two theoretical models', Management Science, 1989, pp. 982-1003

7. European commission directive 2010/40/EU: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:207:0001:0013:EN:PDF [accessed 01 April 2014]

8. European Commission: 'Commission recommendation of 26 May 2008 on safe and efficient in-vehicle information and communication systems; update of the European Statement of Principles on Human-Machine Interface', 2008

9. Auto Alliance: file:///C:/Users/aweare/Downloads/DFTExecutiveSummary2%20(1).pdf [accessed 01 April 2014]

10. NHTSA: 'Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices', 2013. file:///C:/Users/aweare/Downloads/distracted_guidelines-FR_04232013.pdf [accessed 02 April 2014]

11. JAMA: 'Guideline for in-vehicle display systems – Version 3.0'. http://www.umich.edu/~driving/documents/JAMA_guidelines_v30.pdf [accessed 02/April 2014]

12. UNECE: 'Guidelines on establishing requirements for high-priority warning signals', 2011: http://www.unece.org/fileadmin/DAM/trans/doc/2011/wp29/ITS-19-05e.pdf [accessed 02 April 2014]

13. Cotter, S., Stevens, A., Mogilka, A. and Gelau, C.: ‘Development of innovative methodologies to evaluate ITS safety and usability’, HUMANIST TF E. Proceedings of European Conference on Human Centred Design for Intelligent Transport Systems. 3–4 April 2008, Lyon, France. ISBN 978-2-9531712-0-4.

14. IHRA-ITS: ‘Design principles for advanced driver assistance systems‘. http://www.unece.org/fileadmin/DAM/trans/doc/2012/wp29/ITS-20-05e.pdf [accessed 02 April 2014]

15. Burnett, G.E.: ‘On-the-move and in your car: An overview of HCI issues for in-car computing’, International Journal of Mobile Human-Computer Interaction, 2009, 1(1), pp. 60-78

16. Rogers, Y., Sharp, H. and Preece, J.: ‘Interaction design – Beyond human-computer interaction’ (Wiley and Sons: Chichester, 2011, 3rd edn.)

17. Ramney, T.A., Scott Baldwin, G.H., Parmer, E., Domeyer, J., Martin, J. and Mazzae, E.N.: ‘Developing a test to measure distraction potential of in-vehicle information system tasks in production vehicles’. DOT HS 811 463, November 2011

18. ISO/TC 22/SC 13: ‘Road vehicles – Ergonomic aspects of transport information and control systems – Occlusion method to assess visual distraction due to the use of in-vehicle information and communication systems’, International Standard 16673, 2007

19. ISO/TC 22/SC 13: ‘Road vehicles – Ergonomic aspects of transport information and control systems – Simulated lane change test to assess in-vehicle secondary task demand’, International Standard 26022, 2010

20. Pettit, M.A., Burnett, G.E. and Stevens, A.: ‘An extended keystroke level model (KLM) for predicting the visual demand of in-vehicle information systems’, Proceedings of the ACM Computer-Human Interaction (CHI) conference, Held San Jose, May 2007

21. Stevens, A. and Cynk, S.: ‘Checklist for the assessment of in-vehicle information systems’. Transport Research Laboratory, MSC005, 2011