Real-time Assessment of Driver Cognitive Load as a prerequisite for the situation-aware Presentation Toolkit PresTK

Christoph Endres
German Research Center for Artificial Intelligence (DFKI GmbH)
Saarbruecken, Germany
christoph.endres@dfki.de

ABSTRACT

In this paper, we discuss the impact of cognitive load (CL) measuring to the design of In-Vehicle Information Systems (IVIS). Three ways of assessing CL known from literature are discussed in terms of their applicability to non-intrusive in-car applications. A fourth way to estimate CL is presented. We add some thoughts on estimating the complexity of presented information and combine it with the knowledge of the driver's CL and driver's user model to approaches for adapting the system to the driver's current mental state. Furthermore, we shortly introduce our currently developed presentation toolkit PresTK, which takes the discussed factors into consideration.

Categories and Subject Descriptors

H.5 [**Information Interfaces and Applications**]: User Interfaces; H1.2 [**User/Machine Systems**]: Human factors—*complexity measures, performance measures*

General Terms

Theory

Keywords

cognitive load, presentation complexity, automotive information systems

1. INTRODUCTION

The rapid increase of complex information systems in our environment and especially in the car has raised the level of attention required from the user, i.e., the driver. In order to ensure the usefulness of information presented by an in-car Human Machine Interface (HMI), the question whether or not it can be adequately processed by the driver has to be asked. Similar to the specification of the hardware requirements for a certain piece of software, we will have to specify the cognitive requirements of processing a certain piece of information.

Figure 1: The Cognitive Load of the driver can be used for adapting the User Interface

In a situation with high cognitive demand, information presentation could for instance be simplified by removing less relevant information to ensure that the driver registers all important facts.

Although the definitions of Cognitive Load (sometimes also called mental workload or cognitive workload) slightly differ from each other, they are typically similar to Wickens' definition as "the relationship between the cognitive demands of a task and the cognitive resources of the user" [23]. A more detailed definition is given by [3]: "the demands placed on a person's working memory by (a) the main task that she is currently performing, (b) any other task(s) she may be performing concurrently, and (c) distracting aspects of the situation in which she finds herself". [4] provides a survey of alternate definitions.

Traditionally, workload assessment techniques are divided in three groups: subjective measures (questionnaire based, self-reported), performance-based measures, and physiological measures. By widening the scope of assessment beyond actual measuring, we might add a fourth category of deducting cognitive workload from the environment.

Figure 1 shows the connection between the situation/context, the drivers cognitive workload and his driving performance. As we can see, the presentation manager acts as a link in a feedback loop, regulating indirectly the CL by adapting the HMI. This is assisted by presentation complexity estimation. In this paper, we discuss the implications for designing in-car HMIs as a prerequisite for the currently developed presentation toolkit PresTK [9], which tackles the orchestration of mutually independent information sources.

2. COGNITIVE LOAD ASSESSMENT

Knowing the current cognitive state (here: CL) of the driver is very useful for adapting the HMI according to his needs. As a first step, we need to make sure that the measurement reflects reality as close as possible in order to reliably utilize it later on. In this section, we discuss several ways of CL assessment in respect to their usefulness in in-car applications and their feasibility for non-intrusive measuring while driving.

2.1 Subjective measures

A simple and reliable way to assess a subject's workload is self-reporting, assuming that the person is cooperative and capable of introspection and reporting their perceived workload, either directly or by answering questions resulting in a measure. Commonly, questionnaires for self-reporting workload refer to a task already performed. One of the most widely known methods here is the NASA Task Load Index (NASA-TLX). Self-reporting of workload usually covers a single task and cannot be used without extension or modification to report on a complex situation involving several, potentially overlapping tasks. Furthermore, applying questionnaires is an intrusive procedure (adding another task to the subjects working memory) and can only be done after the task being performed. Although some tests are intended to be administered "online" right after performing the task, the test might interfere with the performance in subsequential tasks. Furthermore, none of the online questionnaires are designed for real-time assessment.

It is important to keep in mind that most of these questionnaires are not designed for automotive applications, and not all of them measure the same dimensions–if they are multidimensional at all. Dimensions necessary in the driving context are detailed in [9].

The NASA Task Load Index (NASA-TLX), for instance, which was developed in a "multi-year research effort aimed at empirically isolating the factors that are relevant to subjective experiences of workload" [14] was originally intended for crew complement in the aviation domain. Since its introduction in the mid-eighties, it has spread significantly beyond the original application, focus and language [13]. It is designed as a short questionnaire with 6 questions to be answered on a 21 point scale. The result of the test after a complex evaluation is a multidimensional numerical value on six subscales, only one of them being mental demand.

The Bedford Scale [6] uses a completely different approach. It is a uni-dimensional rating scale designed to "identify operator's spare mental capacity while completing a task". It uses a hierarchical decision tree guiding the user to a rating scale value between one and ten. It is an obvious advantage of the process that in each step of the decision tree the symptoms of having exactly that level of workload are verbally described. This prevents the user from a natural tendency to avoid the extreme values of the scale, even if appropriate.

The Subjective WORkload Dominance (SWORD) technique, as another example, is based on mutual comparison between tasks [22]. The user rates all possible pairs of given tasks in mutual comparison on a scale. A judgment matrix is then calculated based on this data. If this matrix is consistent, relative ratings of each task can be determined.

Figure 2: Connection between situation, cognitive load and driving performance

2.2 Performance based measures

Assuming that an increased CL diminishes human performance, we can use performance measures as an indicator of actual workload. This assumption is backed by the Yerkes-Dodson-Law [24], which is based on an experiment with electric shocks on laboratory mice. Unfortunately, this law has some serious gaps: 1. The methods of calibrating electricity were too crude for exact measurements, 2. the implicit underlying assumption of a linear dependency between stimulus and level of arousal was never validated, and 3. the connection between the behaviour of mice and human beings was just implicitly assumed. Despite these flaws, the use of the Yerkes-Dodson-Law has been established as a valid method [20].

The basic statement is–rephrased for our domain–that the driver's performance is best at a medium level of arousal / workload, i.e., he should neither be bored nor overwhelmed. [12] also examined the impact of cognitive distraction and showed that it has a negative influence on driving performance and safety, especially on the driver's visual behavior. Two approaches of performance measures are feasible in an automotive environment: measuring the driving performance and measuring the reaction time to events such as displayed information or events outside the car.

Driving performance

Recent literature on measuring the drivers CL strongly emphasizes the role of speed and steering wheel angle and their respective change over time. This is very convenient, since this information is easily acquired using the car's CAN-bus. [16] built a prototype to estimate driver distraction in a simulator based on a Fast Fourier Transformation (FFT) of the steering wheel angle. [21] use an artificial neural network (NN) to determine the driver's current level of distraction. Using a three layered Multi-Layer-Perceptron, a single numerical value as the level of distraction ranging from one to five is deducted from four input variables: speed, speed variation, steering wheel angle and steering wheel angle variation. An adaptive system taking driver distraction into consideration was evaluated as being superior in terms of perceived safety and usability to the non-adaptive version. Models based on neural networks have proven successful previously, e.g. [1]. [19] estimates CL complexity using both performance and physiological data in a simulator. As performance measures, the lateral position variation and steering wheel activity is observed. That data is then fed into a radial-basis probabilistic neural network (RBPNN).

Reaction time and time perception

Reaction time is a convenient way of measuring performance. [17] clearly shows a direct impact of driver and situational factors on break reaction time (BRT) and acceleration / deceleration reaction time (ADRT). [5] measured the impact of distraction by mobile phones on the driver's reaction time. Many other examples can be found in literature.

As another interesting aspect, CL seems to directly influence the perception of time. In a user study, [2] measured the difference between time intervals produced by a driver in different situations and compared the mean deviation from the actual time with the CL of the driver measured by other means. Results show a direct connection, i.e., perceived time correlates with actual cognitive workload.

2.3 Physiological measures

Although usually used in medical research and examination for obtaining information of state and performance of major organs, we can also use physiological sensors for obtaining information about the state of the subject. Most suitable for our purpose are obviously parameters which can not consciously be modified. For our purpose, it is important to find a completely non-intrusive method of measuring. Even small intrusion–like placing a sensor on a finger–which is easily accepted in a user study, is unlikely to find acceptance by the driver in every day driving.
Measures known from literature include respiration, skin conductance, temperature, eye movement, pupil diameter, and voice analysis. Only the last three of those can be measured in an unintrusive way, but the analysis of the data can get quite complex. [8] discusses the different methods in detail.

3. COGNITIVE LOAD BY CONTEXT

As we discussed in the previous section, applying traditional CL measuring techniques is not always desirable in our domain. Important features are real-time conduction, immediate availability of results (e.g. results do not have to be entered manually in the system), and unintrusiveness. Table 1 compares the advantages and disadvantages of different approaches.

Measure	Real-Time	Immediate	Intrusive
Subjective	- -	-	- -
Performance	++	+	++
Physiological	++	++	- -

Table 1: Suitability of cognitive load assessment for real time automotive applications is limited.

As shown in Figure 2, current CL might also be estimated using another path, i.e. by assessing the impact of the environment on the driver. Although the context might not be sufficient for an exact estimate of the driver's state, we can safely assume some factors to be influential to his cognitive demands. Driving on the highway or in dense city traffic is probably more demanding than driving on a quiet rural road. Driving at a moderate speed is less stressful than driving at very high speed or being stuck in a traffic jam. Also, environmental conditions such as noise level inside and outside the car can be measured and considered. The cars built-in information systems can keep a history of information presented to the driver, from which we can conduct the cognitive demand. A lot of information flooding the driver in a very short period of time is likely to raise his CL.
[18] used Dynamic Bayesian Networks (DBNs) and data obtained from the car directly to generate a continuous estimate of the drivers load. In a second step, the DBNs were transformed into arithmetic circuits for efficiency reasons, especially considering the usually limited computing power of a vehicle. This concept could be adapted and extended to other information sources in order to increase the quality of the estimate.
In our current research, we examine the impact of visual environmental complexity and speed on the driver's cognitive load in a simulator experiment [11].

Figure 3: A Presentation manager aware of the current cognitive load of the driver can positively influence driving performance

4. ESTIMATING COMPLEXITY

In order to assess system-generated CL, we need to be aware of the impact of system-generated presentations to the driver, i.e. estimate presentation complexity. The approach for answering this question is depending on availability of structured data. We distinguish three different cases:
1. Structured information about presentations is available as a blueprint in the system. In that case, experts can analyze and annotate this information. This enables us to choose the most appropriate presentation type at runtime.
2. When obtaining structured presentations at runtime, we can analyze for instance linguistic complexity, amount of information pieces, font size, complexity of icons and graphics, and other factors. [7] for instance presented a measure for linguistic complexity, which could be used both for analyzing textual on-screen presentation as well as for estimating the complexity of synthesized speech output. Similar measures can be found in literature. [15] provides a very detailed survey and quantitative analysis on the impact of parameters such as font size and contrast on the average glance time of the driver. We propose a layout-based procedure to combine previous research results in [10].
3. We obtain an unstructured presentation in form of an image or an audio file, or both. Chances of making a very good analysis of its complexity in real time are not very good then, but we might be able to give a rough estimate based on formal parameters described in case 2.

5. IMPLICATIONS FOR IVIS

How can we adequately utilize the previously collected information? Figure 3 shows the impact of a presentation manager to the driver's CL. By assessing both information complexity as well as measuring CL, presentations can be modified such that in high demand situations the additional cognitive workload is kept at a minimum. Complex presentations can be avoided or replaced by presentations with a simplified version of the same content, or, in case of low priority, skipped altogether. If complex presentations have to be presented, we can make sure that the time for processing them is sufficiently long. If new and potentially difficult

Figure 4: The presentation toolkit PresTK considers drivers cognitive load and information complexity.

to grasp graphical concepts are used in the HMI, we may consider introducing them in low demand times and only use them as well in high demand times after we can assume the driver's familiarity. Another aspect to be considered is the individual's cognitive capacity. Determining factors are (among others) age, experience and skills.

6. THE PRESTK TOOLKIT

The context of the research presented in this paper is the currently developped presentation toolkit PresTK [9]. Its architecture (see figure 4) reflects both the dynamic nature of the driving environment as well as the double restriction of available resources: There are technical restrictions, e.g., the available space for presenting information is limited, which is followed up by the cognitive limitation of the driver. By both analysing the complexity of the presented information as well as monitoring the current cognitive load of the driver, presented information can be adapted in the scheduling process and be filtered in high demand times.
The toolkit is designed with the automotive domain in mind, but can be used more generally for similar problems as well. By using a component-based structre, we add flexibility to customize several components, the selection of components used, and thus the architecture required in general.

7. CONCLUSIONS

Real-time assessment of the drivers state, especially his CL, is an important factor for adapting IVIS and making the flow of necessary information more efficient without overwhelming the driver. As a foundation, we need a combination of either measuring or estimating CL with an approximate quantification of the complexity of the information presented. The resulting system serves as a regulatory circuit between HMI, driving performance, and CL. Individual need for adaptation may vary among drivers, depending on their cognitive capacity. We discussed the options for necessary building blocks and their suitability for this endeavor in this paper. The concepts presented in this paper provide a part of the foundation for the development of the presentation toolkit PresTK.

8. REFERENCES

[1] D. Akin and B. Akba. A neural network (NN) model to predict intersection crashes based upon driver, vehicle and roadway surface characteristics. *Sci. Res. Essays*, 5(19):2837–2847, 2010.

[2] D. Baldauf, E. Burgard, and M. Wittmann. Time perception as a workload measure in simulated car driving. *Applied ergonomics*, 40(5):929–935, 2009.

[3] A. Berthold and A. Jameson. Interpreting symptoms of cognitive load in speech input. *Courses and Lectures*, pages 235–244, 1999.

[4] B. Cain. A Review of the Mental Workload Literature, 2007.

[5] J. Caird, C. Willness, P. Steel, and C. Scialfa. A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis & Prevention*, 40(4):1282–1293, 2008.

[6] W. Corwin, D. Sandry-Garza, M. Biferno, G. Boucek Jr, and A. Logan. Assessment of Crew Workload Measurement Methods, Techniques and Procedures. Volume 1. Process, Methods and Results. Technical report, DTIC Document, 1989.

[7] V. Demberg and A. Sayeed. Linguistic cognitive load: implications for automotive UIs. In *Adjunct Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2011)*, 2011.

[8] V. Dimitrova-Krause. Physiological Measurement of Driver Stress Induced by Car2X-Based Local Danger Warnings. Master's thesis, Saarland University, 2010.

[9] C. Endres. *PresTK: Situation-Aware Presentation of Messages and Infotainment Content for Drivers.* PhD thesis, Saarland University, 2012. To appear.

[10] C. Endres, M. Feld, and C. Müller. A Layout-based Estimation of Presentation Complexity. In *Adjunct Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2012)*, Portsmouth, New Hampshire, USA, October 2012.

[11] C. Endres, R. Math, and D. Braun. Simulator-based Evaluation on the Impact of Visual Complexity and Speed on DriverŠs Cognitive Load. In *Adjunct Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2012)*, Portsmouth, New Hampshire, USA, October 2012.

[12] J. Harbluk, Y. Noy, and M. Eizenman. The impact of cognitive distraction on driver visual behaviour and vehicle control. Technical report, Transport Canada, 2002.

[13] S. Hart. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 50, pages 904–908, 2006.

[14] S. Hart and L. Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload*, 1:139–183, 1988.

[15] D. Imbeau, W. W. Wierwille, and Y. Beauchamp. *Age, display design and driving perfomance*, chapter 16, pages 339–357. Taylor and Francis London, 1993.

[16] T. Islinger, T. Köhler, and B. Ludwig. Driver Distraction Analysis based on FFT of steering wheel angle. In *Adjunct Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2011)*, Salzburg, Austria, November 2011.

[17] A. Mehmood and S. Easa. Modeling reaction time in car-following behaviour based on human factors. *International Journal of Applied Science, Engineering and Technology*, 5(14):93–101, 2009.

[18] J. Müller. Beanspruchungsschätzung im Automobil mit Bayes'schen Netzen. Master's thesis, Saarland University, 2005.

[19] J. Son and M. Park. Estimating Cognitive Load Complexity Using Performance and Physiological Data in a Driving Simulator. 2011.

[20] M. Staal. Stress, cognition, and human performance: A literature review and conceptual framework. *NASA technical memorandum*, 212824, 2004.

[21] P. Tchankue, J. Wesson, and D. Vogts. The Impact of an Adaptive User Interface on Reducing Driver Distraction. In *Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2011)*, pages 87–94, 2011.

[22] M. Vidulich, G. Ward, and J. Schueren. Using the subjective workload dominance (SWORD) technique for projective workload assessment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 33(6):677–691, 1991.

[23] C. Wickens. Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, 3(2):159–177, 2002.

[24] R. Yerkes and J. Dodson. The relation of strength of stimulus to rapidity of habit-formation. *Journal of comparative neurology and psychology*, 18(5):459–482, 1908.