

Evaluation of the driver's mental workload: a necessity in a perspective of in-vehicle system design for road safety improvement

Annie Pauzié

Published online: 20 May 2014
© Springer-Verlag London 2014

Human factors research and debate related to mental workload have been going on for decades since the 60's (McKenzie et al. 1966) and is still happening (Finomore et al. 2013). The raised issues are: is it useful (“do we need the concept of mental workload or do we have to banish it for good?” Leplat 2002), is it scientifically credible (Dekker et al. 2010), and, in case of a positive answer, how to measure it (Jex 1988).

In the context of the driving task, human factors research is abundant and diversified, aiming at a better understanding of driver behavior and functional capacities in terms of perception, cognition, and motor processes in order to improve road safety (Lee 2008), drivers' mental workload being an important issue to consider in this framework (Dick de Waard 1996). This problematic is even more crucial since the deployment of onboard Intelligent Transport System in the vehicles (Carsten and Nilsson 2001), as human factors have the responsibility to evaluate whether these innovative systems really support the driving task or, on the contrary, lead to distraction and increase mental workload, with potential dramatic consequences in terms of road safety. So, since the beginning of research in this area, the objectives have been to establish methods of assessing fluctuations in mental workload that are sensitive to the various aspects of attentional processing requirements in relation to both external environmental conditions, such as traffic density as well as in-vehicle

conditions, such as competing visual and auditory displays (Pauzié and Amditis 2010).

Workload can be defined as a hypothetical construct that represents the cost incurred by a human operator to achieve a particular level of performance (Hart 1986). Driving performance and the driver's mental workload are both relevant and complementary parameters to consider, knowing that they can vary independently (Yeh and Wickens 1988). Indeed, if the complexity of the task increases, the driver is able to maintain a stable performance to a certain degree, by increasing effort.

Making only performance assessment, such as measurements of vehicle deviation trajectories in relation to visual consultation of in-vehicle system, or quantifying and qualifying driving errors, is not sufficient. These variables reveal only very poor system design inducing high-critical situations. Ambitious of the human factors is firstly, to be subtler in the assessment of the system's usability and, secondly, to get indications about which part of the design would require improvement: visual and/or auditory features, timing of the messages, etc. Assessment of a driver's mental workload will allow obtaining data on the level of effort, if any, to maintain performance, and indications on what was the origin of the costly process (Pauzié 2008).

Furthermore, analysis of observable driver's behavior such as visual strategy in terms of glance duration and glance frequency gives interesting data of the display visual demand (Parkes et al. 1991), and can lead to a standard performance concept, where it is considered that a system is unsafe because it requires too much reading time to be used (Zwahlen and Rockwell 1977). Nevertheless, workload needs to be evaluated in parallel, in order to take also into account the potential cognitive support brought by the displayed information—in comparison with reference situation where there is no system, and so, no visual load

This commentary refers to the original article available at:
doi:10.1007/s10111-014-0275-1.

A. Pauzié (✉)
IFSTTAR, Bron, France
e-mail: annie.pauzie@ifsttar.fr

toward display—such as, for example, the decrease in workload provided at the strategic level of the orientation process in the case of the support of a correctly designed navigation system (Pauzié and Pachiaudi 1997).

For all these reasons, in addition to data coming from performance and behavioral analysis, complementary measurements are necessary in order to identify the driver's workload. In this purpose, workload assessment techniques abound (Cain 2007; Tsang and Wilson 1997). However, subjective ratings are the most commonly used method, as subjective ratings may come closest to tapping the essence of mental workload and would provide the most generally valid and sensitive indicator (Johanssen and Moray 1979) and are also very easy to use from a practical point of view, especially in a context of real road experimentations.

The NASA-TLX is the most popular of these subjective evaluation methods. It assumes that the workload is influenced by several factors, namely *mental demand*, *physical demand*, *temporal demand*, *performance*, *frustration level* and *effort*, with the assumption that some combination of these dimensions is likely to represent the “workload” experienced by most people performing most tasks (Hart and Staveland 1988). Originally, this method was tested and used by the army which considered it as being superior in terms of sensitivity and well accepted by the operator compared with other subjective methods (Hill et al. 1992).

Since its creation, this method has been explosively used in human factors researches in several areas. Hart (2006) investigated this phenomenon 20 years later and noted that simply “Googling” the phrase “NASA-TLX” returned 82,900 citations, 44,000 of which were in English, from diversified areas of investigation and fields of research.

This tool has been also widely used in the context of road safety and design of In-Vehicle Information Systems (IVIS) and Advanced Driver Assistance Systems (ADAS). For example, for IVIS, it allowed evaluating navigation system usability and human-machine interface design (Park and Cha 1998) and, for ADAS, it allowed showing that an Adaptive Cruise Control (ACC) system substantially decreased driver's workload among drivers with lower limb disabilities (Peters 2001). This method allowed assessing efficiency of an adaptive HMI that filters information presentation according to situational requirements in field experiment, in conjunction with objective methods (Piechulla et al. 2003). Data from NASA-TLX ratings allowed also to demonstrate the effect of an increase in driver workload in a study of mobile phoning while driving (Alm and Nilsson 1994). These are some examples of the huge amount of knowledge gathered thanks to this method.

It has to be stressed that mental workload is multidimensional and, among other things, depends upon the type of task. Since NASA-TLX was initially designed for use in

aviation, original factors were adapted to this context. The Driving Activity Load Index (DALI), which is a revised version of the NASA-TLX, has been created specifically for the driving context, with factors defined according to the specific dimensions of this task that could potentially induce workload (Pauzié and Marin-Lamellet 1989). The basic principle is the same as the NASA-TLX, with a scale rating procedure for six pre-defined factors, the weighting procedure being eliminated after several analysis showing the lack of usefulness of this second part of the tool. The six DALI factors are: *Effort of attention*, *Visual demand*, *Auditory demand*, *Temporal demand*, *Interference* and *Situational stress*.

A real road experiment has been conducted in order to define advantages and limits of the DALI method for the evaluation of driver's mental workload (Pauzié and Manzano 2006). As the objective of the experiment was to test the method, the driving conditions were set up to induce on purpose various and diversified levels of workload for the driver. Results showed that the tool reflected correctly what was expected according to the a priori complexity of the driving context, data resulting from the subjective evaluation matching with the characteristics of the various experimental conditions.

The DALI has been used for various research investigations, testing in-vehicle navigation and guidance systems used by young and elderly drivers, and hands-free car phone while driving (Pauzié and Pachiaudi 1997). This method has been also applied for evaluation of IVIS usability (Harvey et al. 2011), and to study the relationship between cognitive workload and driving performance (Gabaude et al. 2012). Tretten (2011), Tretten et al. (2009) analyzed DALI scores in addition to eye-tracking data and interviews to compare four different Head-Up Display (HUD) areas in research investigating IVIS safety use.

In a study comparing navigation based upon a 2D map display and an Augmented Reality (AR) display where virtual objects are superimposed on the real scene, data from DALI indicated that AR navigation was visually and temporally more demanding than the 2D map, and analysis of driver's eye movements confirmed this result, with AR navigation attracting driver's visual attention more frequently than map navigation (Kim and Wohn 2011).

Investigating an innovative mode of driver/system dialogue, where the driver's eye-gaze tracking, in combination with a button on the steering wheel as explicit input, substitute the interaction on the touch screen, allowing then hands to remain on the steering wheel, Kern et al. (2010), using the DALI, concluded that this mode of interaction is more distracting than a touch screen, performance data indicating that it is also a slightly slower mode.

Subjective workload evaluation defined by factors based upon task specificity is efficient and useful for the

researcher; it has been demonstrated with the driving task and the process is ongoing in the context of the riding task. Indeed, research has been conducted recently on the adaptation and implementation for Powered Two Wheeler's of appropriate ADAS/IVIS technologies, renamed Advanced Rider Assistant System (ARAS) and On-Bike Information System (OBIS) that might contribute to significant enhancement of riders' safety. In this framework, it has to be stressed that safety consequences have to be carefully processed, as riding is a very sensitive task with any unexpected distraction or change in rider's dynamic motion potentially leading to loss of control.

In order to evaluate riders' workload, the Riding Activity Load Index (RALI) has been developed after discussion with experts in this area (Pauzié et al. 2009). The aim was to adapt this tool to the specificity of this riding task the same way the NASA-TLX was adapted to the driving one. The first version of the RALI has the following factors: *Visual Demand, Auditory Demand, Temporal Demand, System Interference, Effort Of Attention, Situation Own Coping, Situational Stress, Emotions Handling Vehicle*. Two main factors have been added to the tool, as they seem to be typical from the riding context: "Situation own coping" related to "evaluate the workload induced for coping with the other vehicles and with the complexity of the environment" and "Emotions handling vehicle" related to "evaluate the level of negative emotions linked to the control and the handling of the motorbike". Preliminary tests have been conducted with motorbike manufacturers; additional experimentations would allow improving and validating the method.

1 Conclusion

From an operational perspective in road safety, human factors constructs related to evaluation of drivers' mental workload are essential in prediction. This method allows to identify and to characterize critical contexts inducing high mental workload that could be not detected by performance or behavior, and then making possible to propose concrete and effective improvements in terms of system design, training and adaptation procedures; as Hart and Wickens (1990) pointed out, designers, manufacturers, and operators, who are ultimately interested in system performance, need answers about operator workload.

Subjective measures of mental workload allow "evaluation" rather than "measurement," by establishing relative comparison between situations, and producing global and even "crude" criterion. Nevertheless, it is a powerful and easy to use tool, allowing detecting changes such as resource allocation that would be impossible to identify by direct observation.

Several decades of human factors research in the automotive domain have demonstrated the success of using driver workload evaluation to get reliable and efficient data for avoiding misconception of in-vehicle system design and for allowing better understanding of drivers' behavior, with the purpose of road safety improvement. The methods of workload evaluation have been diverse, with, nevertheless, a high disposition for the subjective evaluation tool, and more specifically the NASA-TLX, with certainly a possible Matthew effect in addition to the willingness for the researchers to be congruent with the other studies conducted in the area.

Of course, it is still relevant and desirable to work on robustness, adaptation and reliability of existing tools, techniques and procedures available to evaluate mental workload; creation of the DALI for the driving context and the RALI for the riding context are examples regarding this objective.

Human constructs research gains by building up knowledge on the representational side of the representational–operational continuum, in addition to this pragmatic operational approach. For example, the neuroergonomics approach deserves to be investigated in the coming decades, bringing innovative knowledge, thanks to the great improvement of physiological recording, to the theoretical background of driver's cognitive processes (Fort et al. 2010).

References

- Alm H, Nilsson L (1994) Changes in driver behaviour as a function of handsfree mobile phones—a simulator study. *Accid Anal Prev* 26(4):441–451
- Cain B (2007) A review of the mental workload literature, RTO-TR-HFM-121-Part-II. Defense Research and Development Canada, Toronto, ON, Canada
- Carsten OMJ, Nilsson L (2001) Safety assessment of driver assistance systems. *Eur J Trans Infrastruct Res* 1(3):225–243. ISSN 1567–7141
- de Waard D (1996) The measurement of drivers' mental workload. The Traffic Research Center VSC, University of Groningen, Haren
- Dekker S, Nyce JM, Van Winsen R, Henriqson E (2010) Epistemological self-confidence in human factors research. *J Cogn Eng Decis Mak* 4:27–38
- Finomore V, Shaw T, Warm J, Matthews G, Boles D (2013) Viewing the workload of vigilance through the lenses of the NASA-TLX and the MRQ, human factors. *J Hum Factors Ergon Soc* (published online 22 April)
- Fort A, Martin R, Jacquet-Andrieu A, Combe-Pangaud C, Foliot G, Daligault S, Delpuech C (2010) Attentional demand and processing of relevant visual information during simulated driving: a MEG study. *Brain Res* 1363:117–127
- Gabaude C, Baracat B, Jallais C, Bonniaud M, Fort A (2012) Cognitive load measurement while driving, HFES Europe Chapter Conference Toulouse

- Hart SG (1986) Theory and measurement of human workload. In: Zeidner J (ed) *Human productivity enhancement*. Praeger, New York, pp 496–555
- Hart SG (2006) NASA-Task Load Index (NASA-TLX); 20 years later, Human factors and ergonomics society annual meeting proceedings, vol 50, pp 904–908
- Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock PA, Meshkati N (eds) *Human mental workload*. North Holland Press, Amsterdam
- Hart SG, Wickens CD (1990) Workload assessment and prediction, chapter 9. In: Booher HR (ed) *An approach to systems integration*. Manprint, Van Nostrand Reinhold, pp 257–296
- Harvey C, Stanton N, Pickering C, McDonald M, Zheng P (2011) A usability evaluation toolkit for In-Vehicle Information Systems (IVISs). *Appl Ergon* 42:563–574
- Hill SG, Iavecchia HP, Byers JC, Bittner AC, Zaldad AL, Christ RE (1992) Comparison of four subjective workload rating scales. *Hum Factors* 34:429–439
- Jex H (1988) Measuring mental workload: problems, progress, and promises. *Adv Psychol* 52:5–39
- Johanssen G, Moray N, Pew R, Rasmussen J, Sanders A, Wickens C (1979) Final report of experimental psychology group. In: Moray N (ed) *Mental workload: its theory and measurement*. Plenum Press, New York, pp 101–116
- Kern D, Mahr A, Castronovo S, Schmidt A, Müller C (2010) Making use of drivers' Glances onto the screen for explicit gaze-based interaction, second international conference on automotive user interfaces and interactive vehicular applications. November 11–12, Pittsburgh, Pennsylvania
- Kim K-H, Wohn K-Y (2011) Effects on productivity and safety of map and augmented reality navigation paradigms. *IEICE Trans Inform Syst* E94D(5):1051–1061
- Lee JD (2008) Fifty years of driving safety research. *Hum Factors* 50(3):521–528
- Leplat J (2002) Éléments pour une histoire de la charge mentale. In: Jourdan M, Theureau J (eds) *Charge mentale: notion floue et vrai problème*. Octarès Éditions, Toulouse, pp 27–40
- McKenzie RE, Buckley EP, Solaranis K (1966) An exploratory study of psychophysiological measurements as indicators of air traffic control sectors workload, Federal Aviation Agency Memorandum Report, 157-524-038, Washington
- Park P, Cha D-W (1998) Comparison of subjective mental workload assessment techniques for the evaluation of in-vehicle navigation system usability, towards the new horizon together. In: Proceedings of the 5th world congress on intelligent transport systems, Seoul, Korea. Paper no 4006
- Parkes AM, Ashby MC, Fairclough SH (1991) The effects of different in-vehicle route information displays on driver behaviour. In: Proceedings of the conference vehicle navigation and information system VNIS 91, pp 61–70
- Pauzié A (2008) A method to assess the driver mental workload: the driving activity load index (DALI). *IET Intell Trans Syst J* 2(4): 315–322
- Pauzié A, Amditis A (2010) Chapter 2: intelligent driver support system functions in cars and their potential consequences on safety. In Ashgate (ed) *Safety of intelligent driver support systems: design, evaluation, and social perspectives*, pp 7–25
- Pauzié A, Manzano J (2006) Subjective evaluation of the driver's workload in real road experiment. European project AIDE, Deliverable 2.2.6
- Pauzié A, Marin-Lamellet C (1989) Analysis of aging drivers behaviors navigating with in-vehicle visual display systems. In: Vehicle navigation and information systems conference, Toronto, pp 61–67
- Pauzié A, Pachiaudi G (1997) Subjective evaluation of the mental workload in the driving context. In: Rothengatter T, Carbonell Vaya E (eds) *Traffic and transport psychology: theory and application*, Pergamon, pp 173–182
- Pauzié A, Gelau C, Aupetit S (2009) Safety evaluation methodology of ITS for riders. International conference models and technologies for intelligent transportation systems, 22–23 June 2009, Rome
- Peters B (2001) Transport for the people with disabilities: barrier-free. *IATSS Res* 25(1)
- Piechulla W, Mayser C, Gehrke H, König W (2003) Reducing drivers mental workload by means of an adaptive man-machine interface. *Trans Res Part F* 6:233–248
- Tretten P (2011) Information design solutions for automotive displays, Focus on HUD, Doctoral Thesis, Luleå University of Technology, Department of Innovation and Design, Division of Industrial Design
- Tretten P, Normark CJ, Gärling A (2009) Where should driver information be placed? A Study on Display Layout. In: HFES 2009 annual meeting, San Antonio, Texas
- Tsang P, Wilson GF (1997) Mental workload. In: Salvendy G (ed) *Handbook of human factors and ergonomics*. Wiley-Interscience Publication, New York, pp 417–449
- Yeh Y, Wickens CD (1988) Dissociation of performance and subjective measures of workload. *Hum Factors* 30(1):111–120
- Zwahlen HT, Rockwell TH (1977) Driver eye movement research: vehicle operations, Human Factors Society Workshop Booklet and Addendum