Transportation Research Part F 32 (2015) 23–34

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

Towards testing auditory–vocal interfaces and detecting distraction while driving: A comparison of eye-movement measures in the assessment of cognitive workload

Michal Niezgoda [a,b,*], Adam Tarnowski [b,1], Mikołaj Kruszewski [a,2], Tomasz Kamiński [a,3]

[a] *Motor Transport Institute, Poland*
[b] *Faculty of Psychology, University of Warsaw, Poland*

A R T I C L E I N F O

Article history:
Received 20 October 2014
Received in revised form 27 April 2015
Accepted 27 April 2015
Available online 17 May 2015

Keywords:
Auditory–vocal interfaces
Distraction detection
Cognitive workload
Driving simulator
n-back
Eye movements

A B S T R A C T

Recently, there has been a growing need among researchers to understand the problem of cognitive workload induced by auditory–verbal–vocal tasks while driving in realistic conditions. This is due to the fact that we need (a) valid methods to evaluate in-vehicle electronic devices using voice control systems and (b) experimental data to build more reliable driver state monitoring systems. In this study, we examined the effects of cognitive workload induced by the delayed digit recall task (*n*-back) while driving. We used a high-fidelity driving simulator and a highway scenario with moderate traffic to study eye movements in realistic driving conditions. This study included 46 participants, and the results indicate that a change in pupil size is most sensitive for measuring changes in cognitive demand in auditory–verbal–vocal tasks. Less sensitive measures included changes in fixation location and blink rate. Fixation durations and the driving performance metrics did not provide sensitive measures of graded levels of cognitive demand.

1. Introduction

Every day, approximately 3500 people are killed in road accidents globally. This fact leads to more than 1.2 million deaths on the world's roads per year. Moreover, millions of people are injured and very often disabled. The World Health Organization (World Health Organization, 2013) estimated that road traffic injuries will become one of the leading causes of death in next decade. Many of these accidents emanate from human error. However, driving itself is an extremely complex activity. It can be evaluated from the perspective of different driving behavior theories and models (Groeger, 2002; Michon, 1985; Reason, Manstead, Stradling, Baxter, & Campbell, 1990). This state is due to the fact that driving outcomes are a result of the interaction of different factors. Nevertheless, researchers were always interested in specific human factors underlying traffic accidents. An analysis using the 100-car Naturalistic Driving Study Data and a data set of 2,000,000 miles conducted by the National Highway Traffic Safety Administration (NHTSA) indicated that inattention and distraction are major factors

---

* Corresponding author at: Motor Transport Institute, Poland. Tel.: +48 22 43 85 127; fax: +48 22 43 85 401.

E-mail addresses: michal.niezgoda@its.waw.pl (M. Niezgoda), adam.tarnowski@psych.uw.edu.pl (A. Tarnowski), mikolaj.kruszewski@its.waw.pl (M. Kruszewski), tomasz.kaminski@its.waw.pl (T. Kamiński).

[1] Tel.: +48 22 55 49 777.

[2] Tel.: +48 22 43 85 405; fax: +48 22 811 09 06.

[3] Tel.: +48 22 43 85 129; fax: +48 22 811 09 06.

http://dx.doi.org/10.1016/j.trf.2015.04.012
1369-8478/© 2015 Elsevier Ltd. All rights reserved.

contributing to near-crash situations and road incidents (Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006). This phenomenon is not a new problem in traffic safety science. However, due to the presence of new types of distractors, there is a growing need among researchers to better understand this cognitive mechanism. In the scope of this study, the theoretical background is closest to the concept of diverting attention toward competing activities and the resulting cognitive workload (Regan, Hallett, & Gordon, 2011).

Cognitive workload itself is a complex construct. Apart from the studies on attention, it has always been central within the working memory domain. This might be due to the fact that there may be one common discrete resource mediating working memory storage and capacity limits of attention (Ester, Vogel, & Awh, 2012). However, in applied psychology and human factors studies, cognitive workload usually refers to the concept of imposing demands on humans' limited mental resources. There is a certain limit of resources that, when exceeded, task realization performance significantly decreases. Cognitive workload is often studied in one of two paradigms, single-task demand or dual-task demand. Multiple tasks are related to the concept of multiple resource theory that deals with time-sharing efficiency when performing concurrent tasks (Wickens, 2002; Wickens, Toplak, & Wiesenthal, 2008). This theory indicates different dimensions of information processing related to the sensory modalities, codes, responses, and stages underlying dual-task performance.

Drivers are heavily exposed to multitasking. When we are driving, we perform concurrent tasks. The act of driving is mainly processed in the visual–spatial–manual pathway, but it can be also affected by engaging in the secondary audi tory–verbal–vocal task, such as phone conversations (Ho & Spence, 2008). There is also a second use case for this type of information processing. The evolution of human–machine interfaces toward auditory interfaces in recent years has affected the design of in-vehicle devices. In some studies, these interfaces are labeled as easier and more satisfying to use than visual interfaces (Sodnik, Dicke, Tomazič, & Billinghurst, 2008). However, it also triggered questions about road safety implications. Lee, Caven, Haake, and Brown (2001) reported that speech-based systems simulating the work of e-mail significantly increases reaction time and introduces a major cognitive workload. Concerns over workload induced by voice-based interactions are also reflected in study by Reimer and Mehler (2013), where authors concluded that these interactions can also lead to increased visual demand.

There are several ways to measure cognitive workload, both objective and subjective. One of the most popular among researchers is still subjective assessment conducted using the NASA task load index (NASA-TLX) (Hart & Staveland, 1988) and the Subjective Workload Assessment Technique (SWAT) (Reid & Nygren, 1988). Their popularity is due to their low cost, non-intrusiveness, ease-of-use, known sensitivity, and validity (Zhang & Luximon, 2005). On the other hand, there exists a group of objective techniques aimed at measuring cognitive workload, which seem to be more applicable to the real-time monitoring of the driving activity. The first group of objective measurement techniques can be classified as performance indicators, and includes steering performance metrics, lane keeping, speed management, vehicle following, response time, and steering grip (Ariën et al., 2013; Minin, Benedetto, Parotid, Re, & Thesauri, 2012). The second group includes physiological and neurological measures. In the context of multitasking while driving, heart rate and skin conductance showed consistent patterns in detecting an increased workload (Mehler, Reimer, Coughlin, & Dusek, 2009). Other good indicators of workload in this group can be amplitude of P3b in electroencephalography (EEG) analysis (Lei, Welke, & Roetting, 2009). Other activation patterns in the human brain that correspond to the working memory load and visual attention load can be detected when using functional magnetic resonance imaging (fMRI) (Tomasi, Chang, Caparelli, & Ernst, 2007). However, it should be noted that this technique is not easily applicable to driving studies but not impossible (Hunga et al., 2014). Developments in the automotive industry and implementations in other related industries, such as those that involve earth-moving equipment, eye-tracking can be considered a very promising direction for real-time monitoring of the operator's state.

Within the eye-tracking domain, researchers use different variables to measure cognitive workload, such as changes in the diameter of the pupil, blink rate, blink duration, saccadic movements, dwell rate, fixation durations, fixations rate, and fixation locations. These measures are defined by standardization bodies and researchers e.g. International Standards Organization (2002), National Highway Traffic Safety Administration (2012) and Holmqvist (2011). The size of the pupil is one of the most commonly used workload indicators, and it has long been recognized in a variety of cognitive tasks, such as arithmetic problems and linguistic tasks. It is usually calculated as percentage change in pupil diameter due to the pupil diameter variability among humans. Apart from the cognitive activity, pupil size is affected by lighting conditions. In stable lighting conditions, pupil size increases with an increasing cognitive workload (Iqbal, Zheng, & Bailey, 2004; Tsai, Viirre, Strychacz, Chase, & Jung, 2007). For instance, the pupil diameter in highly demanding arithmetic tasks can change by more than 20%, but the percentage of this change varies between different types of tasks and cognitive efforts (Holmqvist, 2011). Results by Recarte and Nunes (2000) indicate that pupil size can change during the spatial-imagery and verbal tasks performed while driving. However, Demberg, Sayeed, Castronovo, and Müller (2013) conclude that there are more reliable and more sensitive pupil measures than just simple changes in pupil size, for example, the Index of Cognitive Activity (ICA) (Marshall, 2000, 2007). This complex measurement system calculates rapid and small pupil dilations, and has proved to be a good indicator of detecting cognitive workload in linguistic tasks and digit-span tasks in simulated driving conditions (Demberg et al., 2013; Schwalm, Keinath, & Zimmer, 2008).

Blink rate and blink duration are highly discussed measures of cognitive workload. Both measures seem to be correlated with the number of errors in visual tasks (Van Orden, Jung, & Makeig, 2000). Tsai et al. (2007) found that blink frequency increases in the dual-task paradigm among drivers when the second task is an auditory task. In the same study, blink duration was not statistically significant. Other experimental results indicate that blink duration and blink frequency are highly dependent on the visual demand imposed by a single task or two concurrent tasks (Recarte, Pérez, Conchillo, & Nunes, 2008).