

Document Version

Final published version

Licence

CC BY-ND

Citation (APA)

Marquart, G., Cabrall, C., & de Winter, J. (2015). Review of eye-related measures of drivers' mental workload. In T. Ahrm, W. Karwowski, & D. Schmorow (Eds.), *Procedia Manufacturing - 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015* (Vol. 3, pp. 2854-2861). (Procedia Manufacturing; Vol. 3). Elsevier. <https://doi.org/10.1016/j.promfg.2015.07.783>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the
Affiliated Conferences, AHFE 2015

Review of eye-related measures of drivers' mental workload

Gerhard Marquart, Christopher Cabrall*, Joost de Winter

Department of BioMechanical Engineering, Delft University of Technology, Mekelweg 2, 2628 CD, The Netherlands

Abstract

The assessment of mental workload could be helpful to road safety especially if developments of vehicle automation will increasingly place drivers into roles of supervisory control. With the rapidly decreasing size and increasing resolution of cameras as well as exponential computational power gains, remote eye measurements are growing in popularity as non-obtrusive and non-distracting tools for assessing driver workload. This review summarizes literature on the relation between eye measurement parameters and drivers' mental workload. Various eye activity measures including blinks, fixations, and saccades have previously researched and confirmed as useful estimates of a driver's mental workload. Additionally, recent studies in pupillometry have shown promise for real-time prediction and assessment of driver mental workload after effects of illumination are accounted for. Specifically, workload increases were found to be indicated by increases in blink latency, PERCLOS, fixation duration, pupil dilation, and ICA; by decreases in blink duration and gaze variability; and with mixed results regarding blink rate. Given such a range of measures available, we recommend using multiple assessment methods to increase validity and robustness in driver assessment.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of AHFE Conference

Keywords: Eye tracking; Mental workload; Driving

1. Introduction

Alarming proportions of human errors such as misperception, information processing errors, and poor decision making are frequently identified as causes of accidents [1]. The assessment of drivers' mental workload could be helpful in improving driving tasks to reduce the number of accidents, which are largely attributed to the drivers themselves. As automation is more rapidly being developed than in years before, poorly designed automated systems

* Corresponding author. Tel.: +31-015-27-85608.
E-mail address: c.d.d.cabrall@tudelft.nl

could lead to situations characterized by the phrase “99% boredom, 1% terror”. This phrase refers to supervisory control situations in which an operator experiences boredom due to mental underload for most of the time and terror due to mental overload in case of emergencies. Physiological measures appear to be an attractive workload assessment approach as these can be obtained without mediation by subjective response or transformation through (secondary) performance manifestation. However, wiring drivers with sensors and electrodes is rarely feasible nor desirable and furthermore many bodily measurements lack temporal resolution required for real-time applications. Consequently, without operator-sensor contact requirements and in an era of rapidly decreasing size and increasing resolution of cameras and accelerating gains in computational power, remote eye measurements are growing in popularity as non-obtrusive and non-distracting techniques for assessing driver workload.

Driving is not only a physical (e.g., applying force on steering wheel and pedals) but also a visual and a mental task. The eyes of a driver are indispensable in performing visual tasks such as scanning the road, communicating with other road users, and monitoring in-vehicle devices. Mental tasks are important during driving, and include such factors as understanding vehicle dynamics, making situation-dependent decisions, and judging time/space relationships. Visual and mental tasks are closely linked to each other (e.g., scanning a crossroad while simultaneously judging the time/space relationships of other road users). It is therefore of no surprise that many researchers use or study eye-related parameters to assess drivers’ mental workload [2-6].

Probably the first extensive literature review on the topic of measuring drivers’ mental workload was conducted by De Waard [4]. Since then, much research has been done in this field, resulting in new insights and the development of new measurement techniques. Borghini et al. [7] reviewed the literature related to the neurophysiological signal measurements of mental workload, fatigue, and drowsiness of aircraft pilots and car drivers. However, this particular review was limited to the electrooculography (EOG) measurement method and did not cover other eye-related techniques. The aim of this review is to summarize the results of recently conducted studies about the relation between eye measurement parameters and driver mental workload. The review proceeds with definitions before continuing with summary and discussion of the related findings.

2. Definitions

In addition to many different possible definitions of mental workload, many different eye measurements might be investigated in relation to it. For the purpose of this review, definitions follow first for mental workload and the common eye measurements of blinks, fixations, and pupillometry.

The general agreement according to Recarte et al. [8] is that “mental workload is the result of an interaction between task demands and human characteristics, that is, separately, neither the task properties nor the human operator characteristics can explain mental workload”. Similarly, Hart and Staveland [9] defined workload as a hypothetical construct representing the cost incurred by human operators to achieve a specific level of performance (i.e., workload is not the same as objective task demands). Mental workload in driving specifically was defined by Boer [10] as the effort required to maintain the driving state within a subjective safety zone. In this definition, the subjective safety zone depends on the driver’s needs in terms of costs and benefits.

There are three fundamental types of eye blinks: reflexive, voluntary, and endogenous. Endogenous eye blinks are distinguished from other blinks by the absence of an identifiable eliciting stimulus and can reflect the mental workload induced by task demands [11]. Commonly used blink parameters include blink rate, blink duration, and blink latency (i.e., the length of time between the occurrence of task relevant information and the subsequent blink initiation). Another blink-related parameter commonly used is PERCLOS, defined as the percentage of time that the eyelid covers 80% or more of the pupil [12].

Over the years, many parameters characterising eye fixations have been studied in relation with driver’s mental workload, including number of fixations, fixation duration, number of saccades, saccadic duration, saccadic amplitude, peak of saccadic velocity, and gaze distribution [6], [13-18]. Because fixation duration and spatial gaze distribution were frequently reported parameters between those listed above, the current review focuses on these.

Pupillometry is the measurement of the pupil diameter, which is considered to be a valid indicator of mental workload [2], [19-22]. Small and involuntary fluctuations of the pupil diameter up to 0.5 mm (cf. 2 to 8 mm for responses to changes of light) have been identified as a reaction to a cognitive processing task and are called the task-evoked pupillary response (TEPR). Examples of such mental tasks are short and long-term memory access,

mental arithmetic, sentence comprehension, vigilance, and visual and auditory perception [2]. It is important to note that since TEPRs do not occur reliably, multiple TEPRs must be recorded and averaged to obtain a reliable estimate of an individual's mental workload. Most pupillometry studies make use of TEPR-averaging methods. Another method to measure mental workload using pupillometry is the index of cognitive activity (ICA) [23]. The ICA is a signal processing technique that uses the different reflex properties to separate the effects of illumination and mental workload on the pupil's size.

3. Results

3.1. Drivers' mental workload

In the literature, researchers usually distinguish three categories of mental workload measures: performance, subjective, and physiological measures [24]. Performance measures are used to determine how well an operator, or driver in this case, is performing one or more tasks. If driving a vehicle would be the primary task, then a secondary task could be used to determine the driver's remaining mental capacity, assuming that both tasks use the same mental capacity pool. Secondary task performance measures can be either superficially or closely related to the driving task itself (see [25] for a conceptualization of the driver's task). Performance-based measures of drivers mental workload can thus be seen to be challenged by issues of adaptive resource strategies of drivers who may drive in different manners and in varying relations with other tasks.

Subjective measures capture the driver's own assessment of mental workload. Frequently used assessment tools like the NASA Task Load Index (NASA-TLX, [9]), the Rating Scale Mental Effort (RSME, [26]) or other types of questionnaires [25] are administered after an experiment and so capture the mental workload experienced across a task overall. Tattersall and Foord [27] adopted an instantaneous self-assessment (ISA) method, originally designed for air traffic control, and used for on-line measurements. Statistically significant correlations were found in this study between the ISA scale and other mental workload ratings, such as heart-rate variability. However, the results in this study also indicated that the ISA method interferes with the driver's performance, making it less useful.

Physiological measures record the driver's physiological state during a drive. Examples of these measures are: eye activity (e.g., pupil dilation, blink rate, and eye fixations), head motion, brain activity, heart rate, blood pressure, muscular activity, body movement and posture, endocrine response, and galvanic skin response [28]. Besides mental workload, these physiological measures are also influenced by other aspects of the driver's mental and physical state (e.g., physical fatigue and circadian rhythm) and by environmental variables (e.g., illumination and temperature) [29]. Most physiological measurements also require specialized equipment and technical expertise [5] however due to miniaturisation of the measurement equipment many physiological measures can be collected continuously and relatively unobtrusively [4]. The focus in the remainder of this review is on eye-related physiological measures of driver mental workload.

3.2. Blinks

Blink rate has been investigated in a series of driver and workload studies with mixed results attributable to the distinction between mental and visual workload. Kramer (1990) stated in his review of physiological measures of mental workload that the results related to blink rate were mixed, with sometimes increasing rates and sometimes decreasing rates depending on the visual demands [5]. His recommendation was that more research was required before blink rate could be used as a measure of mental workload. Subsequently, in a real car-driving study published in 1998 [30], mental workload was captured with a number of physiological measures, including blink rate. The results in this study showed that the eye blink rate decreased as the curves of the road became sharper. Ten years later Recarte et al. [8] analysed the eye blink rate in single- and dual-task (cognitive task plus visual search) conditions and found that the blink rate increased for all cognitive tasks (listening, talking, and calculating) when compared to the control condition. The researchers also found a decrease in eye blink rate for more visually demanding tasks when compared to less visually demanding tasks. Recarte et al. therefore concluded that

“according to blink rate, visual and mental workload produce opposite effects: Blink inhibition for higher visual demand and increased blink rate for higher mental workload” [8].

This theory might help explain the mixed results obtained in earlier studies where visual and mental workload were not clearly differentiated, as in Heger’s [30] study where the results showed a blink rate decrease as the curves become sharper. The driver’s task in this experiment might be more visually demanding (scanning the curved road ahead) than mentally demanding (controlling the vehicle), which could explain the decrease in blink rate.

Blink rate has also been investigated in relation with highly automated driving, a system that automates a vehicle’s motion in the lateral as well as the longitudinal direction. Systems like these are typically designed, among other things, to decrease a driver’s mental workload. It is remarkable, however, that several studies found a blink rate increase during highly automated driving compared to normal driving [31-33].

Blink duration has been shown to decrease with increases in mental workload. The studies mentioned in Kramer’s [5] review all found shorter blink durations for increasing task demands (both mental and visual). In flight simulator studies researchers also reported a decrease in blink duration as visual workload increases [34-35]. In a recent driving simulator study [36], the authors examined the effects of in-vehicle information systems (IVIS) on eye blinks while performing a lane change task (LCT) during driving. The IVIS task was used as a secondary task and required visual perception and manual response. The authors assumed that this task used the same mental capacity pool and would therefore increase the visual and mental processing demands. The results showed a blink duration inhibition for the dual-task (LCT and IVIS task) compared with the single-task condition (LCT). According to the authors, this blink duration inhibition may occur to avoid visual information loss. This thought follows well with Kramer’s own earlier conclusions that blinks are postponed until sufficient visual information is obtained [5].

Regarding the presentation of task relevant information and the latency until the next subsequent blink, Kramer provided multiple examples of studies in his review in which the latency of blinks increases with increasing mental task demands [5]. The author stated that it is likely that eye blinks are postponed until sufficient visual information is extracted from the environment to perform the task as well as possible. Similar results of increasing blink latency for increasing mental task demands were found in studies conducted by [37-38].

PERCLOS values have been found to positively correlate with increasing subjective sleepiness, performance decrements, and the number of lapses in a visual reaction time task [39-42]. The results obtained by [42] showed that PERCLOS correlated even better with performance lapses in the reaction task than the participants’ own ratings of sleepiness. Halverson et al. compared the use of several eye metrics, including PERCLOS, for classifying mental workload [43]. The authors knew that PERCLOS was a popular well-verified measure of fatigue but thought it was surprising that it had not yet been studied in relation with mental workload. Ten extensively trained individuals participated in his study by performing visual searches on simulated radar images while driving an autonomous vehicle in a simulated environment. The results showed that PERCLOS can be used to distinguish between two different levels of workload. Halverson et al. also stated that this measure and the pupil diameter complement each other for the assessment of mental workload [43].

3.3. Fixations

Dwell time and fixation duration are extensively used measures and are generally believed to increase with increasing mental task demands [6], [44-47]. In the years after, fixation duration has also been investigated in a series of studies in relation with driver hazard perception as described in Underwood et al.’s literature review [15]. All the reviewed studies showed increased fixation durations during hazardous moments, indicating increased mental workload. Underwood et al. therefore stated that “Since long fixation durations are typically associated with high processing load, it makes sense to think that during these hazards viewers are spending longer extracting information from their point of gaze” [15].

Previous research demonstrated that increased attentional demand (obtained by increasing both the visual and mental demands) produces attentional focus narrowing, also called spatial gaze concentration [6], [46], [48-49]. Recarte and Nunes [6] were of the first to investigate the effect of mental instead of visual tasks on eye fixation parameters during driving in real traffic. Their results showed lower gaze variability in both horizontal and vertical spatial gaze direction even for verbal (only mentally demanding) secondary task conditions. Similar results of spatial gaze concentrations due to added mental workload were found in later studies ([14-15], [17]). In the on-road study

conducted by [14], drivers' gaze distributions were found to be significantly smaller while performing additional auditory-only secondary tasks, especially during the most cognitively difficult tasks. It is interesting to note that in this study the drivers seemed to change their visual search behaviour (smaller gaze distributions) before vehicle control suffered. Lastly, another study by Recartes and Nunes [13] aimed to figure out in real traffic whether visual-detection impairment and spatial gaze concentration (induced by mental tasks) equally affects the entire visual field (general interference) or the peripheral areas in particular (tunnel vision). The results showed that this impairment was independent of target eccentricity and the authors therefore concluded that a general interference effect was produced instead of a tunnel vision effect.

3.4. Pupillometry

In two on-road driving studies, secondary cognitive tasks were used to induce mental workload during driving and recorded pupil dilation [6] [13]. Their results showed a consistently increased pupil diameter during their spatial imagery and verbal secondary tasks. Palinko et al. [29] used a remote eye tracker in a driving simulator study to estimate driver's mental workload. Again, pupil diameter was used as a reliable way of measuring the driver's mental workload, especially in simulated driving environments. Palinko et al. also introduced a new measuring variant of the pupil diameter called the mean pupil diameter change rate (MPDCR) [29]. They suggested that this new measure might be particularly useful when rapid changes in the mental workload should be detected.

Subsequent studies explored the effects of mental workload and illumination on the pupil diameter in more detail. For example, with participants instructed to fixate their gaze on one of three differently coloured vehicles (black, grey, and white trucks), Palinko and Kun showed that it is possible to separate the effects of lighting and mental workload on the pupil diameter using a light reflex predictor, which tries to predict the reflex of the pupil to illumination [50]. The prediction is then subtracted from the pupil signal, leaving an estimation of the TEPR. Driver's mental workload has also been investigated by Kun et al. in relation with in-vehicle spoken dialogues. Participants were instructed to follow a vehicle at a comfortable distance [51]. During the ride, the drivers played a series of word games with a remote conversant with the intention to vary the mental workload. The driver's TEPR was analysed for two points in time: 1) just before the first contribution of the remote conversant and 2) just before the first contribution of the driver's response. It was found that the pupil diameter was significantly larger at point 2 than at point 1 in 69% of all word games. Several possible explanations are given by the authors for this fairly weak effect, including the possibility that in some cases the TEPR may have been masked by the pupillary light reflex. So once again, researchers warn against the negative effects of light reflexes on pupil dilation.

Additionally, the ICA method has been used in a number of studies and has shown to be a promising method of measuring driver's mental workload. Schwalm was one of the first to show that the ICA may be a reliable method to identify driver's mental demands in both highly controlled driving simulators as well as in real environments [52]. Demberg et al. used the ICA method to estimate mental workload during simulated driving and found increased ICA values resulting from more difficult speech comprehension clauses [53]. The ICA method was also used by Dlugosch in a static simulator study using a head-mounted eye tracker [54]. Participants were asked to do several secondary tasks, including three conversation tasks and a surrogate reference (visual search and manual response) task (SuRT). Significantly higher resultant ICA values were found for the most demanding secondary task condition compared to all other conditions, while all secondary tasks had higher ICA values than for driving only.

4. Discussion

The aim of this review is to summarize the results of recently conducted studies about the relation between eye measurement parameters and driver mental workload. As a result of the review, various eye measures have been observed in the literature as confirmed estimates of a driver's mental workload with the following relations (Table 1). Specifically, increases in workload were found to be indicated by increases in blink latency, PERCLOS, fixation duration, pupil dilation, and ICA; by decreases in blink duration and gaze variability; and with mixed results regarding blink rate.

Table 1. Relation of eye-related physiological measures and drivers' mental workload.

Measure		Mental workload
Blinks	Rate	+ / -
	Duration	-
	Latency	+
	PERCLOS	+
Fixations	Duration	+
	Gaze variability	-
Pupillometry	Pupil dilation	+
	ICA	+

As described in this review, there are four common blink-related parameters that are used in the literature to assess a driver's mental workload: blink rate, blink duration, blink latency, and PERCLOS. Blink rate has shown to be sensitive to visual and mental workload, and fatigue. Both types of workload produce opposite effects according to Recarte et al. ("Blink inhibition for higher visual demand and increased blink rate for higher mental workload" [8]). However, several studies found a blink rate increase during highly automated driving compared to normal driving [31-33]. This could suggest that highly automated driving relieves a driver more from the visual tasks than from the mental tasks. Blink latency and blink duration have shown to be more straightforward or less nuanced indicators of workload (i.e., non-discriminatory between a mental or visual workload component). The results seem to be consistent: blink duration decreases and blink latency increases for both increases in mental and visual processing demands. Kramer (1990) argued that blink durations decrease and latencies increase to extract task-relevant information from the environment as efficient as possible [5].

Although PERCLOS is mainly used in the literature as a measure of fatigue, Halverson et al. demonstrated that it could also be used to distinguish between two different levels of workload [43]. The authors encouraged researchers to combine different measures when assessing mental workload to increase accuracy and robustness. This suggestion is sensible, because mental workload is a broad concept and might not best be described by a one-dimensional construct, especially in driving where visual demands play a major role as well. Additionally, individual differences like age, risk-proneness, and driving experience have shown to affect drivers' mental workload, which makes assessing it even more challenging. In this context of measurement, Brookhuis et. al (2003) distinguish between absolute and relative criteria [55].

Fixation duration is found to be sensitive to workload and is generally believed to increase with increasing visual and mental processing demands [6], [47]. A series of studies later reconfirmed this finding by proving that fixation duration also increases when approaching potential hazards during driving [16]. Several other studies have demonstrated that visual and mental processing demands produce visual-detection impairment and attentional focus narrowing [6], [14-15], [17], [46], [48-49]. A recent study showed that horizontal gaze dispersion is more sensitive to drivers' mental workload than a weighted average between horizontal and vertical gaze dispersion [56].

Pupillometry has also been used in driving studies to measure drivers' mental workload. In highly controllable environments, such as driving simulators, the results are consistent and show that the pupil diameter is sensitive to changes in mental workload. However, it is believed that pupil dilation, unlike blink rate, does not discriminate between mental and visual workload. Palinko et al. introduced a new measuring variant of the pupil diameter called the mean pupil diameter change rate. They suggested that this new measure might be particularly useful when rapid changes in the mental workload have to be detected [29].

The index of cognitive activity (ICA; patented by [23]) is a promising method that can be used to identify drivers' mental workload in both highly controlled driving simulators and in real environments due to its ability to account for illumination effects [52-54], [57]. This index is calculated through an algorithm that uses pupillometry, but is unfortunately patented and therefore cannot easily be used in research.

In summary, there are many different ways to measure a driver's mental workload but no single method works in all cases, since driving is a multi-dimensional task and all measures have specific drawbacks. The focus for future research should be on combining multiple assessment methods to increase validity and robustness, which is

particularly important when applied in real cars. Advanced measures, like the ICA approach, might play a major role in this ongoing development towards a complete understanding of the driver's psychophysiological state.

Acknowledgements

The authors are involved in the Marie Curie ITN: HFAuto (PITN-GA-2013-605817).

References

- [1] J. Treat, N. Tumbas, S. McDonald, D. Shinar, R. Hume, R. Mayer, R. Stanisfer and R. Castellan. Tri-level study of the causes of traffic accidents. *Report No. DOT-HS-034-3-535-77 (TAC)*, 1977.
- [2] J. Beatty and B. Lucero-Wagoner. The pupillary system. In J. Cacioppo, L. Tassinary & G. Bertson (Eds.), *Handbook of Psychophysiology*, Cambridge: Cambridge University Press, 2000.
- [3] K. Brookhuis and D. De Waard. Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis and Prevention*, 42, pp. 898-903, 2010.
- [4] D. De Waard. The measurement of drivers' mental workload (*Doctoral dissertation*). University of Groningen, Haren, Netherlands. 1996.
- [5] A. Kramer. Physiological metrics of mental workload: A review of recent progress. In D. Damos (Ed.), *Multiple-task Performance*, London: Taylor & Francis, pp. 279-328, 1990.
- [6] M. Recarte and L. Nunes. Effects of verbal and spatial-imagery task on eye fixations while driving. *Journal of Experimental Psychology: Applied*, 6, pp. 31-43, 2000.
- [7] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience and Biobehavioral Reviews*, 44, pp. 58-75, 2012.
- [8] M. Recarte, E. Perez, A. Conchillo and L. Nunes. Mental workload and visual impairment: Differences between pupil, blink, and subjective rating. *The Spanish Journal of Psychology*, 11(2), pp. 374-385, 2008.
- [9] S. Hart and L. Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: North Holland Press, pp. 139-183, 1988.
- [10] E. Boer. Behavioral entropy as a measure of driving performance. *Proceedings of the 3rd International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, Rockport, Maine, pp. 225-229, 2005.
- [11] J. Stern, L. Walrath, and R. Goldstein. The endogenous eyeblink. *Psychophysiology*, 21, pp. 23-33, 1984.
- [12] W. Wierwille, L. Ellsworth, S. Wreggit, R. Fairbanks, and C. Kirn. Research on vehicle-based driver status/performance monitoring: development, validation, and refinement of algorithms for detection of driver drowsiness. *Final Report. DOT-HS-808-247*, Washington, D.C., National Highway Traffic Safety Administration.
- [13] M. Recarte and L. Nunes. Mental workload while driving: Effects on visual search, discrimination, and decision making. *Journal of Experimental Psychology: Applied*, 9, pp. 119-137, 2003.
- [14] B. Reimer. Impact of cognitive task complexity on drivers' visual tunneling. *Transportation Research Record: Journal of the Transportation Research Board*, 2138, pp. 13-19, 2009.
- [15] B. Reimer, B. Mehler, Y. Wang, and J. Coughlin. The impact of systematic variation of cognitive demand on drivers' visual attention across multiple age groups. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 54(24), pp. 2052-2055, 2010.
- [16] G. Underwood, D. Crundall and P. Chapman. Driving simulator validation with hazard perception. *Transportation Research Part F: Traffic Psychology and Behavior*, 14, pp. 435-446, 2011.
- [17] T. Victor, J. Harbluk and J. Engstrom. Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F: Traffic Psychology and Behavior*, 8(2), pp. 167-190, 2005.
- [18] M. Young and J. Mahfoud. Driven to distraction: Determining the effects of roadside advertising on driver attention. *Ergonomics Research Group Report*, Brunel University, Uxbridge, England, 2008.
- [19] W. Heinrich. Die aufmerksamkeit und die funktion der sinnesorgane. *Zeitschrift fur Psychologie und Physiologie der Sinnesorgane*, 9, pp. 342-388, 1896.
- [20] E. Hess and J. Polt. Pupil size relation to mental activity during simple problem solving. *Science*, 143, pp. 1190-1192, 1964.
- [21] J. Beatty. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Physiological Bulletin*, 91 pp. 276-292, 1982.
- [22] D. Kahneman. *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall, 1973.
- [23] S. Marshall. Method and Apparatus for Eye Tracking and Monitoring Pupil Dilation to Evaluate Cognitive Activity. *U.S. patent no. 6,090,051*. 2000.
- [24] R. O'Donnel and F. Eggemeier. Workload assessment methodology. In K. Boff, L. Kaufman & J. Thomas (Eds.), *Handbook of Perception and Human Performance. Volume II, Cognitive Processes and Performance*, 1986.
- [25] J.C.F. de Winter, R. Happee, M. Martens, N. Stanton. Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behavior*, 27, Part B, pp. 196-217, Nov. 2014.
- [26] F. Zijlstra and L. Van Doorn. The construction of a scale to measure perceived effort. *Department of Philosophy and Social Sciences*, Delft University of Technology, Delft, Netherlands, 1985.

- [27] A. Tattersall and P. Foord. An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics*, 39(5), pp. 740-748, 1996.
- [28] P. Desmond. Fatigue and stress in driving performance (*Doctoral dissertation*). University of Dundee, Scotland, UK, 1997.
- [29] O. Palinko, A. Kun, A. Shyrokov and P. Heeman. Estimating cognitive load using remote eye tracking in a driving simulator. *Proceedings of the Eye Tracking Research and Applications Conference*, 2010.
- [30] R. Heger. Driving behavior and driver mental workload as criteria of highway geometric design quality. *Transportation Research Circular*, Washington, DC, Transportation Research Board pp. 43:1 – 43:10, 1998.
- [31] D. Cha. Driver workload comparisons among road sections of automated highway systems. *Proceedings of the Society of Automotive Engineers World Congress*, Detroit, MI. Technical Paper 2003-010119, 2003.
- [32] D. Dambock, T. Weibgerber, M. Kienle and K. Benglar. Requirements for cooperative vehicle guidance. *Proceedings of the 16th International IEEE Annual Congerence on Intelligent Transportation Systems, ITSC*, The Hague, Netherlands, pp. 1656-1661, 2013.
- [33] N. Merat, A. Jamson, F. Lai and O. Carsten. HAD, secondary task performance and driver state. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 547, pp. 762-771, 2012.
- [34] U. Ahlstrom and F. Friedman-Berg. Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics*, 36, pp. 623-636, 2006.
- [35] J. Veltman and A. Gaillard. Physiological indices of workload in simulated flight task. *Biological Psychology*, 42(3), pp. 323-342, 1996.
- [36] S. Benedetto, M. Pedrotti, L. Minin, T. Baccino, A. Re, and R. Montanari. Driver workload and eye blink duration. *Transportation Research Part F: Traffic Psychology and Behavior*, 14(3), pp. 199-208, 2011.
- [37] F. Eggemeier, D. Biers, C. Wickens, A. Andre, D. Vreuls, E. Billman and J. Schuereen. Performance assessment and workload evaluation systems: analysis of candidate measure. *Technical Report No HSD-TR-90-023*. Brooks Air Force Base, Armstrong Aerospace Medical Research Laboratory, TX, 1990.
- [38] M. Carmody. Current issues in the measurement of military aircrew performance: a consideration of the relationship between available metrics and operational concerns. *Air Vehicle and Crew Systems Technology Department*, Naval Air Warfare center: Aircraft Division, Warminster, PA, 1994.
- [39] D. Dinges and R. Grace. PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance. *TechBrief NHTSA*, Publication no. FHWA-MCRT-98-006, 1998.
- [40] F. Friedrichs and B. Yang. Camera-based drowsiness reference for driver state classification under real driving conditions. *Proceedings of the IEEE Intelligent Vehicles Symposium*, San Diego, CA, pp. 101-106, 2010.
- [41] K. Kozak, R. Curry, J. Greenberg, B. Artz, M. Blommer and L. Cathy. Leading indicators of drowsiness in simulated driving. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, 49, pp. 1917-1921, 2005.
- [42] M. Mallis. Evaluation of techniques for drowsiness detection: Experiment on performance based validation of fatigue tracking technologies (*Doctoral dissertation*). Drexel University, Philadelphia, PA, 1999.
- [43] T. Halverson, J. Estep, J. Christensen and J. Monnin. Classifying workload with eye movements in a complex task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 56, 168-172, 2012.
- [44] M. Maltz and D. Shinar. Eye movements of younger and older drivers. *Human Factors*, 41, pp. 15-25, 1999.
- [45] J. May, R. Kennedy, M. Williams, W. Dunlap and J. Brannan. Eye movement indices of mental workload. *Acta Psychologica*, 75, 1990.
- [46] T. Miura. Active function of eye-movement and useful field of view in a realistic setting. In R. Groner, G. D'Ydewalle, & R. Parham (Eds.), *From Eye to Mind: Information Acquisition in Perception, Search, and Reading*, Amsterdam, Netherlands, pp. 119-127, 1990.
- [47] K. Rayner and R. Morris. Do eye-movements reflect higher order processes in reading? In R. Groner, G. D'Ydewalle, & R. Parham (Eds.), *From Eye to Mind: Information Acquisition in Perception, Search, and Reading*, Amsterdam, Netherlands, pp. 191-204, 1990.
- [48] T. Miura. Coping with situational demands: A study of eye-movements and peripheral vision performance. In A. Gale, M. Freeman, C. Haslegrave, P. Smith & S. Taylor (Eds.), *Vision in Vehicles*, Amsterdam, Netherlands, pp. 126-137, 1986.
- [49] R. Radach. Eye guidance and visual information processing: reading, visual search, picture perception and driving. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception*, Oxford, England: Elsevier, pp. 1-27, 1998.
- [50] O. Palinko and A. Kun. Exploring the influence of light and cognitive load on pupil diameter in driving simulator studies. *Proceedings of the 6th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, 2011.
- [51] A. Kun, O. Palinko, Z. Menenica and P. Heeman. On the feasibility of using pupil diameter to estimate cognitive load changes for in-vehicle spoken dialogues. *Proceedings of Interspeech*, pp. 3766-3770, 2013.
- [52] M. Schwalm. Pupillometrie als Methode zur Erfassung mentaler Beanspruchungen im automotiven Kontext (*Doctoral dissertation*). Saarbrücken: Universität Saarbrücken, 2009.
- [53] V. Demberg, A. Sayeed, A. Mahr and C. Muller. Measuring linguistically-induced cognitive load during driving using the ConTRe task. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Automotive UI – 13*, Eindhoven, Netherlands, pp. 176-183, 2013.
- [54] C. Dlugosch, A. Conti and K. Benglar. Driver distraction through conversation measured with pupillometry. *Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, Bolton Landing, New York, 2013.
- [55] K. Brookhuis, D. Waard, and S. Fairclough. Criteria for driver impairment. *Ergonomics*, 46, pp. 433-445, 2003.
- [56] Y. Wang, B. Reimer, J. Dobres and B. Mehler. The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive demand. *Transportation Research Part F: Traffic Psychology and Behavior*, 26, pp. 227-237, 2014.
- [57] V. Demberg. Pupillometry: The Index of Cognitive Activity in a Dual-Task Study. *Proceedings of the 35th annual meeting of the Cognitive Science Society*, cogsci-13, 2013.