Visual Fixations as a Rapid Indicator of Hazard Perception

Boris M. VELICHKOVSKY[1*], Alexandra ROTHERT[1], Darius MINIOTAS[2], Sascha M. DORNHÖFER[1], Markus JOOS[1], Sebastian PANNASCH[1]

[1]*Applied Cognitive Research Unit / Psychology III*
Dresden University of Technology, D-01062 Dresden, Germany

[2]*TAUCHI, Dept. of Computer and Information Sciences*
University of Tampere, FIN-33014 Tampere, Finland

Abstract. Eye movements can provide valuable information about driver's attention and the course of behaviour in hazardous situations. The major goal of this study was to investigate the dynamic switchover between preattentive and attentive processing taking place during a simulated driving task. A PC-based driving simulator was used to create the dynamic environment in conjunction with the Eye-Link™ head-mounted system, which recorded eye movements. Two types of hazards were presented to the subjects: traffic lights and a pedestrian facing the street. The analysis was focused mainly on the situations containing immediate hazards, i.e. a red traffic light and a pedestrian crossing the street. There was a marked increase in fixation duration around the time of emergence of an immediate hazard. This phasic increase was time-invariant. It is therefore possible to track the switching from one level of processing to another with reference to analysis of driver's visual fixations. The study is concluded by a discussion of attentional conditions where overlooked or not sufficiently processed hazards do not lead to the appropriate braking reaction on the part of the driver.

1. Introduction

1.1 Preattentive scanning versus attentive processing

Hazard perception refers to the identification of dangerous traffic situations as they arise. It therefore demonstrates a skill which is of principal relevance for the driving activity (e.g. [1]). According to Nagayama [2], more than 50 percent of all collisions in road traffic arise from missing or delayed hazard perception. In a study of road safety, Treat et al. [3] found that human error was the sole cause in 57% of all accidents and was a contributing factor in over 90%. Only 2.4% were due solely to mechanical fault and only 4.7% were caused by environmental factors. This failure of drivers to perceive road-traffic hazards is often due to the fact that the driver failed to attend, because his/her mental resources were focussed elsewhere – a phenomenon known as "inattentional blindness" [4].

As in the case of many other practically relevant situations, events leading to human errors in traffic scenarios simply are too fast to be addressed by the conventional psychophysiological methods. This is why we analyse in the present article, as well as in several earlier publications (see [5], [6]), behavioral data on eye movements and parametrs of visual fixations, which seem to be the most direct manifestations of visual attention.

[*] Corresponding author. Tel.: +49-351-463-34221; fax: +49-351-463-37741.
E-mail address: velich@applied-cognition.org

First of all, we should consider some elementary facts about cognitive neurophysiology of human information processing.

Following a long tradition in psychology and neuroscience, visual information processing can be described in terms of a two-level model (e.g. [7], [8], [9]). At the first, preattentive level, objects are dynamically localised in the 3D environment – hence the other name for the level, which is 'ambient'. These objects, or rather 'blobs', are identified at the second level, which is variously named as focal, or attentive because it is thought to form an informational bottleneck by operating only on a few objects at a time. For instance, the driver attends to a small subset of road-traffic information while ignoring the rest (e.g. [10]). Once there is attentive processing, eye movements are under endogenous control, which generally leads to longer fixations. The multilevel view is supported by visual search studies of Pomplun [11]. Accordingly, fixations seem to be under control of hierarchical mechanisms – during a visual search-and-compare task, two phases could be distinguished. In the first phase, fixation duration (on average, 150 to 250 ms) is mainly a function of spatial density and configuration. In the second phase, just prior to the solution, fixation duration increases to 500 ms and is no longer controlled by salient physical features, but instead by the complexity of decision, i.e. by higher-level mechanisms responsible for hypothesis generation and testing.

The two-level model certainly is a simplification as there are levels 'above' the form-oriented focal stage. In particular, conceptually-driven (semantic), and self-referential (or metacognitive) processes characterise these mechanisms that residue in the phylogenetically new frontal structures of brain [12]. Furthermore, training and expertise lead to the automatisation of skills, so that with time their components can be processed at lower levels. Nevertheless, the two-level model is a useful first approximation to consider visual fixations from the multilevel perspective. One can expect that higher levels of encoding of visual material may be correlated with longer fixations. Indeed, levels of encoding in visual memory tasks could be isolated by the analysis of fixations [13]. The question is whether these methods of analysis can be applied to driving activity. The purpose of the article is to examine the fixation parameters as indicators of the level of attentional control in hazardous driving situations.

1.2 Distribution of fixation durations and saccades in dynamic scenes

The first step in investigating eye movements from the perspective of attention deployment is to look at the distribution of fixation times (e.g. [14]). A further task is to relate these fixation durations to the amplitude of saccadic movements. In experiments with perception of naturalistic pictures, we found three segments of fixation duration that are related to distinct combinations of the amplitude of saccades. The shortest fixations, below 90 ms, often result from the large ones and are followed by very small saccades. They can be interpreted as stops on the way of correcting the eye position. The next two segments, though not as sharply differentiated in the overall distribution, are of more interest for our analysis. Fixations from 90 to about 140 ms produce large saccades of more than 4°, beyond the parafoveal region of retina. In other words, these saccades aim at targets seen as blobs not as individualised objects – a strong case for preattentive processing. Fixations longer then 140-200 ms seem to be related to focal processing: they initiate saccades mainly within the parafoveal region where objects are relatively easily seen and continuously attended [15].

Our previous driving simulation study revealed that a similar pattern also emerges in interaction with motion flow scenes [16]. The main difference was only that the segments of preattentive and attentive processing in this case were shifted towards longer fixation durations. The reason for this is that static fixations have been transformed into dynamic ones, by including a smooth-pursuit component. Thereafter in the article, the word 'fixation' will be used in the broader sense of 'dynamic fixation'.

2. Methods

The results discussed here are derived from a study of eye movements and hazard perception during a driving task realised on a PC-based driving simulator. The 'SIRCA' driving simulator was developed and adapted to our eye tracking system in cooperation with the ARTEC group (Institute of Robotics, University of Valencia). A video beamer with SXGA-resolution (JVC DLA G11) projected a fictive urban scenario with two to four lanes streets and numerous intersections onto a 1.5 x 2 meter screen. The subject was seated at a distance of 3.5 meters resulting in a view of approximately 24° vertically and 32° horizontally. Eye movements were recorded with the EyeLink™ head-mounted system with an accuracy of better than 1° and 250 Hz sampling rate.

Twelve male subjects, aged 24 to 36 years, took part in the study. All of them were holders of a driving licence (for at least 7 years) with a high driving experience ranging from 70 000 to 200 000 km. Subjects were instructed to drive through the urban environment with a recommended speed of 50 km/h or less, following traffic rules and normally keeping the car on the right lane. Before the experiment, subjects were asked to complete a test drive on the simulator in order to get accustomed to it as well as to the eye tracking equipment. The duration of an experimental drive session was ca. 40 min, whereby the first 20 min were used as an extended training phase. The stable virtual environment of the second, test-phase was always the same, while all the dynamic aspects were randomised. The eye tracker was shortly re-calibrated after 10 min of drive. The session was repeated for all subjects 5 times – one time on each of 5 consecutive weeks.

During the experiment, potential and immediate hazards were presented on the road. Subjects had to react to the hazards in an appropriate way. Figure 1 shows typical screenshots with examples of hazards as seen by the subjects. Potential hazards, which appeared on average approximately every 50 seconds, were defined as situations demanding attentive monitoring to objects that could turn into immediate hazards (what really happened in the 40% of the cases). As potential hazards, either a green traffic light or a pedestrian facing the road as if intended to cross it (Figure 1A) could be presented well before the time subjects were to make a decision for braking. Immediate hazards were situations where the driver had to react immediately in order to prevent an accident.

The immediate hazards, which always evolved out of potential hazards, could be either a red traffic light (Figure 1B) implemented at road crossings only, or a pedestrian crossing the street in front of the subject's car. These hazards always appeared at a distance of 25 m. At the defined maximal speed of 50 km/h, the distance corresponded to the braking distance just sufficient to react to the hazard on time. In addition to the hazardous events, a variety of traffic-related and traffic-unrelated neutral events were presented, such as cars in other lanes or even in the subject's one but at a sufficient distance from them, as well as pedestrians walking on their respective sidewalks independently of the traffic condition.

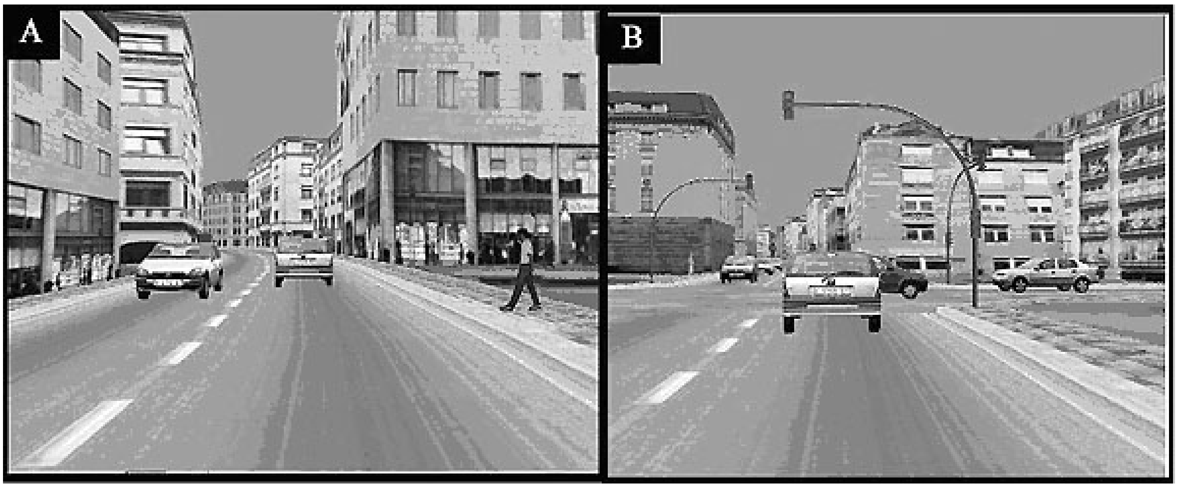

Figure 1. Sample screenshots taken from the experiment. (A) A potential hazard: a pedestrian facing the street; (B) An immediate hazard: the red traffic light at a crossing.

3. Results

3.1 Saccade amplitude and fixation duration

The overall distribution of fixation durations is presented in Figure 2A. It is log-normal, positive skewed with the mode (204 ms) below the mean (400 ms). This finding is consistent with that obtained in our previous driving simulation study [16].

Figure 2B shows the combination of preceding and following saccadic amplitudes. Skipping the few corrective fixations of the shortest segment (<90 ms), one can see the pattern resembling that observed in [16]. Fixations within the segment from about 90 to almost 300 ms are generally related to large-scale ambient exploration, whereas longer fixations demonstrate rather their involvement into a more piecemeal focal processing. We can conclude that on the basis of the overall distribution, fixations sub-serving preattentive processing can be coarsely dissociated from those in service of attentive elaboration. The rule of thumb here is the following: pre-modal and modal fixations are preattentive, mean and post-mean fixations are attentive.

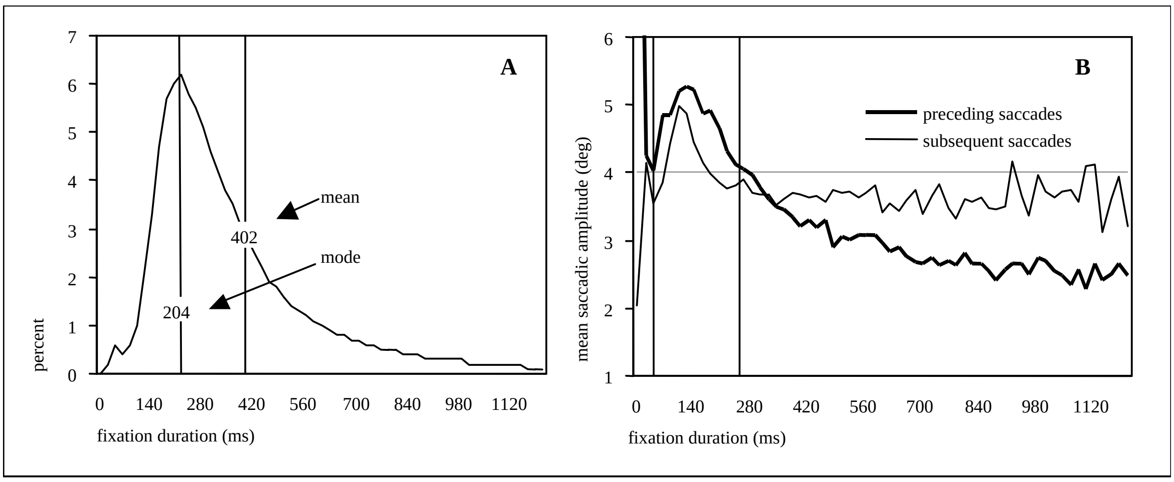

Figure 2: (A) The frequency distribution of fixations (131 654 counts) from the experiment; (B) Distributions of preceding and following saccade amplitudes across the range of fixation durations.

3.2 Immediate hazards

What happens to visual fixations when a critical event occurs? The previous analysis is too crude to answer this question. Since fixation durations may change instantaneously, from one fixation to the next, the duration of the fixation that actually "detects" an immediate hazard may be seen on the background of several preceding fixations. We therefore selected the moment of appearance of the immediate hazard as a starting point, and analysed fixations occurring around this time. Figure 3 shows the average fixation durations plotted over the fixation number relative to an immediate hazard: four fixations before the fixation that was actually hit by the event (numbered as "0"), and the next five fixations. Data for the 5 consecutive drives are presented to see if the reaction declined over time.

The data in Figure 3 reveals that for both types of an immediate hazard a sudden increase in fixation duration upon their emergence is obvious ($F_{1,11}$=6.987, $p<0.001$ and $F_{1,11}$=9.444, $p<0.001$ for red light and walking pedestrian, respectively). The reaction is not only strong but also surprisingly stable – for example, in difference to oculomotor reactions to meaningless distractors, which usually habituate after several repetitions of the event [17]. Analysis of variance supported this impression by providing evidence that the phasic increase of fixation duration does not change over time ($F_{4,11}$=2.347, $p>0.05$ and $F_{4,11}$=1.962, $p>0.05$ for red light and walking pedestrian, respectively).

3.3 Potential and immediate hazards versus base-line

For our further analysis the base-line data on fixation duration can be useful. To find their parameters, we randomly selected sequences of 10 fixations from the driving episodes that had neutral but no hazardous (potential or immediate) events. The procedure was repeated 30 times for every of 12 subjects at any of 5 test-drives resulting in 1800 datasets in total. The resulting distribution of the mean fixation durations was log-normal, left-skewed with the mean of 412 ms and the median of 387 ms. A percentile analysis demonstrated that 5% and 95% of all estimates corresponded to fixation durations of 295 and 600 ms, respectively. These base-line data are visualised in all figures below – the black line illustrates the median and the dashed lines show 45% deviation below or above the base-line median.

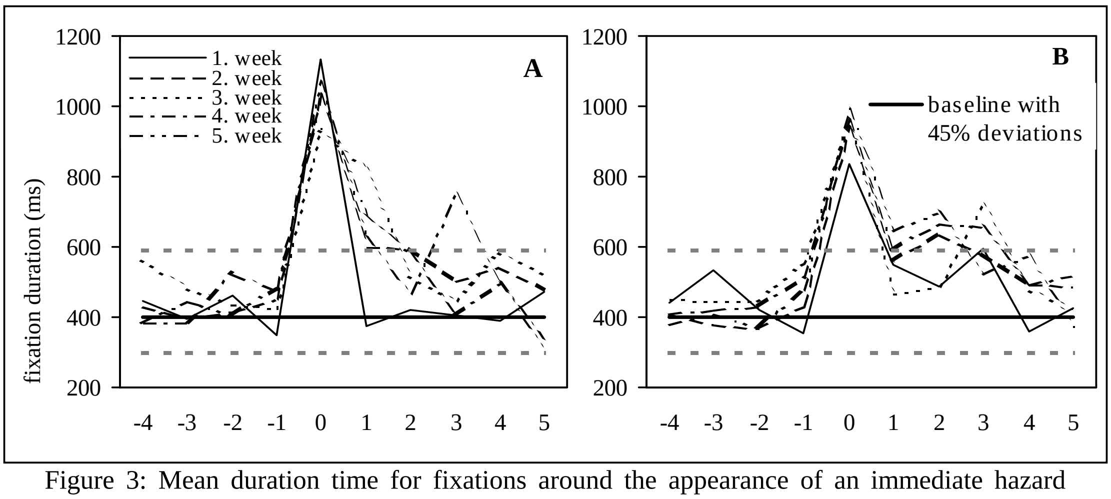

Figure 3: Mean duration time for fixations around the appearance of an immediate hazard (overall 375 counts; "0" corresponds to the fixation at the moment of the critical event), separately for 5 consecutive drive trials: A. Red light; B. Walking pedestrian.

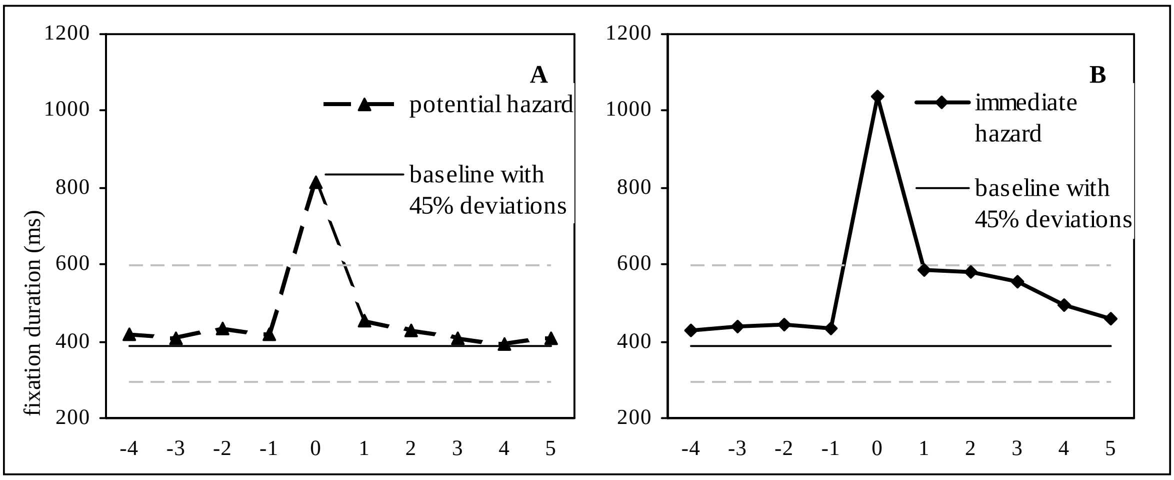

Figure 4. (A) Mean fixation durations around the “point of no return” for potential hazards (“0”corresponds to fixation when the driver is at the distance of about 25 meters to hazard); (B) Mean fixation duration for immediate hazards.

What happens to experienced drivers’ visual behaviour when they approach a potential hazard? We analysed fixation duration when potential hazards reached the distance of 25 m (i.e. the distance where they could be transformed into immediate hazards). Assuming the maximal speed of 50 km/h, this is the last chance for drivers to decide if they start braking or drive through. Therefore, subjects need to attentively monitor the potential hazards at this point of time. As can be seen in Figure 4A, there is an increase in fixation duration when potential hazards were at 25 m distance from the driver ($F_{1,11}$= 7.914, $p < 0.001$). This increase was not significantly different from that of immediate hazards ($F_{1,11}$ = 1.166 , $p > 0.25$). Again, we found no habituation over time ($F_{4,11}$= 0.441, $p > 0.5$).

Another interesting effect can be seen if we compare data in Figures 4A and 4B. In comparison to potential hazards, the three subsequent fixations after the critical event in the case of immediate hazards are prolonged ($F_{1,11}$= 6.749, $p < 0.02$). The same short-term ‘freezing’ of fixations has been already reported in the driving simulation study mentioned above [18]. It is not completely clear whether this effect reflects the driver’s attempt to take up additional visual feedback during their braking, or it is purely an inhibitory by-product of the voluntary motor action per se [19].

Finally, an analysis of the spatial location of fixations is also of interest. The gaze position analysis revealed, for instance, that 81% of the immediate hazards were already viewed foveally (eccentricity of the object location, <1°) or, at least, parafoveally (eccentricity <4°) at the moment when they were presented. Furthermore, at the distance of 25 m, 74% of all potential hazards were too in the foveal or parafoveal view. Thus, if a potential hazard reaches the critical point for braking decision, most of the experienced drivers already attentively track it and are prepared for an adequate action.

3.4 Correct versus accident behaviour in the face of an immediate hazard

To examine the specifics of driver’s fixation behaviour associated with driving through the red light, the situations in which the subjects performed correctly (i.e., stopped for red) were compared with those few cases (N = 12) where accident drives through the crossing occured (Figure 5A and 5B, respectively).

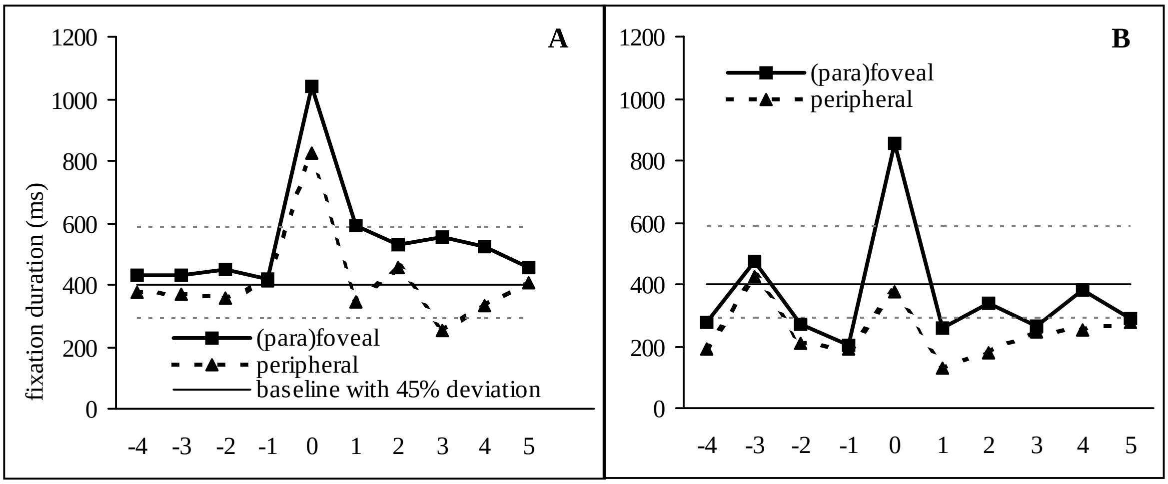

Figure 5: (A) Mean fixation durations around the appearance of red light in the cases of successful braking, categorized by eccentricity. (B) Mean fixation durations for the passing through red light (“0” corresponds to the fixation at the moment of switching to red), categorized by eccentricity.

Gaze position was taken into account for this analysis, too. In the case of a successful braking, our data show an increase in fixation duration at the “point of no return”. This reaction was especially strong when the red light was viewed foveally and parafoveally ($F_{1,11}$=6.473, p <0.001), compared to ($F_{1,11}$=2.213, p<0.05) in the peripheral condition. In the (para)foveal condition the increase of fixation duration with respect to base-line was 637 ms, so that the mean fixation duration was 1038 ms. Nevertheless, an increase in fixation duration was observed even when the subjects did not look at the red light.

A similar trend in driver’s fixation behaviour can be observed for the situations where the role of an immediate hazard was performed by a walking pedestrian (Figure 6). Once again, there is a noticeable increase in fixation duration at the critical point. For those cases where braking action was duly performed by the subjects, the effect was significant for the (para)foveal condition ($F_{1,11}$=9.714, p<0.001). For the peripheral viewing condition, however, the effect was not shown to be significant ($F_{1,11}$=1.214, p>0.05).

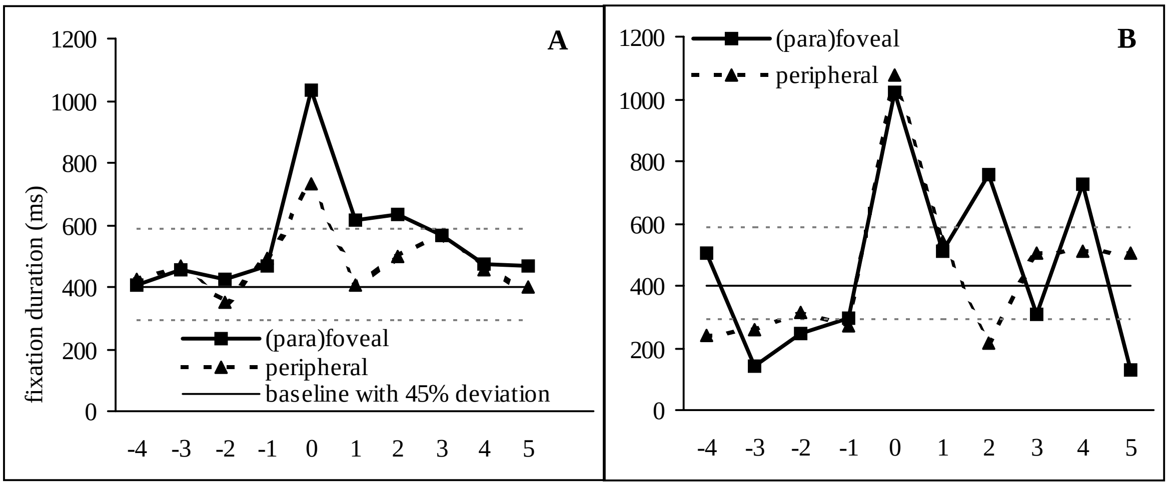

Figure 6: (A) Mean fixation durations around the appearance of a walking pedestrian in the cases of successful braking, categorized by eccentricity. (B) Mean fixation durations for driving over a pedestrian (“0” corresponds to the fixation at the moment the pedestrian starts crossing the street), categorized by eccentricity.

Of particular interest, of course, are the very rare cases (<1%) where the subjects failed to react properly and drove through the crossing on the red light (N = 12, Figure 5B), or over a walking pedestrian (N = 9, Figure 6B). The errors cannot be explained solely by a shortage of decision time as the driving speed in all the cases was slightly below the normative limit of 50 km/h. Due to a small number of observations no regular statistical analysis can be provided, but the duration of fixations at the critical event is clearly above the 95% base-line threshold, especially in the more upsetting cases where pedestrians were involved in the accidents.

The only real difference compared to the data reflecting correct driver response is that the durations of two or three consecutive fixations preceding the crucial point are at or below the 5% base-line threshold, and this observation holds true for both types of the hazard under current investigation. From the point of view of the two-level model of visual perception, these fixations of about 200 ms are preattentive. Thus, the reason for the lack of attention to immediate hazards by some of our subjects in those several cases may simply be that the critical event hit them in their preattentive mode of processing.

4. General Discussion

In the recent years, a variety of physiological functions and indices have been considered with respect to their diagnostic value for the hazard perception in driving tasks (see [20]). The oculomotor parameters are only one source of this information, but also one that can be highly promising for the needed 'express-diagnostics' of driver's attention. As noticed in [21], eye movements provide important insights for understanding the driving tasks and for developing training strategies and accident countermeasures. The results of the present experiment demonstrate that eye tracking is useful to analyse not only the direction but also the level of attention. In particular, visual fixation responses to hazardous events were strong, reliable and fast with an increase of 100 and more percent in fixation duration upon detection of a hazardous event. There was also remarkable stability of these reactions over time, even when the laboratory situation as well as the virtual environment became more and more familiar to the subjects.

Although the phasic response of visual fixations to hazards was strong and always highly significant, this parameter alone was not sufficient to predict what subjects exactly did. For instance, the response to the potential hazards (that were ultimately ignored) was nearly the same as that registered for the immediate ones, which were acknowledged by braking and stopping the car. When subjects erroneously drove through a crossing on the red light, a significant prolongation of the visual fixation was found as well (see Figure 5B). This is the reason why a more theoretical approach to the analysis of driver's oculomotor behaviour is needed.

We attempted to provide such a theoretical framework in terms of the two-level model. The distinction between preattentive scanning and attentive elaboration in visual search seems to be supported by the data. The two-stage search process can be recognized by inspection of fixation behaviour. On a preliminary basis, we propose that fixations subserving preattentive scanning are dissociable from those in service of attentive elaboration on the basis of their duration as well as saccade amplitudes. Of practical interest is that this information can be obtained without taking into account the exact spatial location of eye. By categorization of the level of processing (preattentive/ambient versus focal/attentive) and by monitoring phasic changes in fixation duration, a more reliable diagnosis for the course of driver's behaviour can be made.

At the present time, it is hardly possible to propose an elaborate version of a cognitive architecture that would describe perceptual and cognitive aspects of driver's behaviour. Anderson's ACT-R theory is a good candidate for the role (see [5]), and we certainly can hope that further investigations will bridge the gap between this computational approach and more neurophysiologically oriented work, as it is reflected in the two-level model of visual processing. Indeed, besides empirical data, the last model is mainly supported by the neurophysiological data on the existence of dorsal and ventral systems of human brain (e.g. [22]). In the course of its evolution, the two-level model should include higher-order mechanisms corresponding to the control structure of frontal lobes with their metacognitive and self-referencial types of processing (see [13], [12]). These control functions are at the heart of the global cognitive models like Anderson's ACT-R or Newell's Soar. At least then a synthesis will be possible. The road is long but the intermediate goals are known and within reach.

5. Conclusion

A high-speed monitoring of visual fixations can be used to evaluate driver's attentional state and interpret variation in fixation duration and saccadic amplitude in a simulated driving task. In particular, driver's reaction to potential and immediate hazards demonstrates that eye movement data can be used for an express-diagnostics of behaviour. An important precondition of this development, however, is a reliable differentiation between the ambient and focal processing modes as well as between the higher – semantic and metacognitive - levels of information processing).

Acknowledgements

This work was supported by the BMW AG, Munich (Project "Hazard's Cognition"). We would like to thank Marcos Fernandez Marin for his help in programming the SIRCA-simulator. We also thank Pieter J.A. Unema for discussing the results.

References

[1] J. Crick and F. P. McKenna, Hazard perception. Can it be trained? Behavioural research in road safety II. Proceedings of a seminar at Manchester University, Crowthorne, Berkshire, 1991, pp. 100-107.

[2] Y. Nagayama, Role of visual perception in driving, *IATSS Research* **2** (1978) 64-73.

[3] J. R. Treat, N. S. Tumbas, S. T. McDonald, D, Shinar, R. D. Hume, R. E. Mayer, R. L. Stanisfer and N. J. Castellan, Tri-level study of the causes of traffic accidents. Report No. DOT-HS-034-3-535-77 (TAC), 1977.

[4] A. Mack and I. Rock, Inattentional blindness. MIT Press, Cambridge, 1998.

[5] D. D. Salvucci and A. Liu, The time course of a lane change: Driver control and eye movement behaviour. *Transportation Research*, Part F **5** (2002) 110-119.

[6] B. M. Velichkovsky, A. Rothert, M. Kopf, S. M. Dornhöfer and M. Joos, Towards an express-diagnostics for level of processing and hazard perception. *Transportation Research*, Part F **5** (2002) 145-156.

[7] J. E. Hoffman, Stages of processing in visual search and attention. In B. H. Challis and B. M. Velichkovsky (eds.), Stratification in cognition and consciousness. John Benjamins, Amsterdam/Philadelphia, 1999, pp. 43-71.

[8] C. Trevarthen, Two visual systems in primates, *Psychologische Forschung* **31** (1969) 321-337.

[9] B. M. Velichkovsky, Visual cognition and its spatial-temporal context. In F. Klix, J. Hoffmann and E. van der Meer (eds.), Cognitive research in psychology. North Holland, Amsterdam, 1982, pp.30-48.

[10] B. M. Velichkovsky, S. M. Dornhöfer, M. Kopf, J. Helmert and M. Joos, Change detection and occlusion modes in road-traffic scenarios. *Transportation Research*, Part F **5** (2002) 99-109.

[11] M. Pomplun, Analysis and models of eye movements in comparative visual search. Cuvillier, Göttingen, 1998.
[12] B. M. Velichkovsky, Level of processing. Validating the concept. In M. Naveh-Benjamin, M. Moscovitch and H. L. Roediger, III (eds.), Perspectives on human memory and cognitive aging: Essays in honour of Fergus Craik. Psychology Press, Philadelphia, 2002, pp. 38-73.
[13] B. M. Velichkovsky, From levels of processing to stratification of cognition. In B. H. Challis and B. M. Velichkovsky (eds.), Stratification in cognition and consciousness. John Benjamins, Amsterdam/Philadelphia, 1999, pp. 203-226.
[14] N. Moray, Designing for attention. In A. Baddeley and L. Weiskrantz (eds.), Attention, selection, awareness and control. Clarendon Press, Oxford, 1993, pp. 67-81.
[15] S. Pannasch, S. M. Dornhöfer, P. J. A. Unema, S. Zapf, and B. M. Velichkovsky, Separating ambient and focal fixations, Proceedings of the 11th European Conference on Eye Movements, Turku, Finland, 2001, pp. 22-25.
[16] B. M. Velichkovsky, S. M. Dornhöfer, S. Pannasch and P. J. A. Unema, Visual fixations and level of attentional processing. Proceedings of the International Conference Eye Tracking Research and Applications, Palm Beach Gardens, Florida, 2000, pp. 79-85.
[17] B. M. Velichkovsky and S. Pannasch, The fastest visual reaction adapts to environment, not retinal location, *Behavioral and Brain Sciences* **24**-5 (2001) 1008-1009.
[18] S. Pannasch, S. M. Dornhöfer, P. J. A. Unema, and B. M. Velichkovsky, The omnipresent prolongation of visual fixations, *Vision Research* **41** (2001) 3345-3351.
[19] T. M. Bujakas and N. D. Linde, Inhibitory effect on saccadic eye movements in the process of activity. In A. N. Leontiev (ed.), Perception and activity. Moscow University Press, Moscow, 1974, pp. 68-86.
[20] J. A. Groeger, Understanding driving. Psychology Press, Hove, 2000.
[21] J. F. Chapman and G. Underwood, Visual search of dynamic scenes. In G. Underwood (ed.), Eye guidance in reading and scene perception. Elsevier, Oxford, 1998, pp. 269-294.
[22] A. D. Millner and M. A. Goodale, The visual brain in action. Oxford University Press, New York, 1995.