Visual Fixations and Level of Attentional Processing

Boris M. Velichkovsky[1], Sascha M. Dornhoefer[1], Sebastian Pannasch[1] and Pieter J.A.Unema[2]

[1]Dresden University of Technology

[2]University of Maastricht

1 INTRODUCTION

The analysis of visual fixations can provide rich information about the user's attention. We will show that it is possible to distinguish between preattentive scanning and cognitive elaboration using parameters of fixation duration. It is, however, not without difficulty to interpret fixation durations solely on the basis of averages. First of all, fixation durations show a log-normal rather than a normal distribution, with a more or less strong leftward skew. Van Zandt & Ratcliff [13] stipulated that frequency distributions contain information that is not always represented by the usual first order statistics. This is particularly the case when the distribution is multimodal, but also when experimental factors provoke changes that do not affect the whole distribution, which is usually the case with fixation durations. Moreover, as will be shown below, the same mean fixation duration may result from quite different distributions. Furthermore, the mode is very stable, and can only be altered under few conditions. Finally, the distribution of fixation durations is not unimodal. In most of our experiments, a distinct bimodality can be observed with the second mode lying around 60 ms. These fixations are often ignored, possibly because of the low temporal resolution of the most recording devices. We found these "express-fixations" to be selectively sensitive to task conditions.

1.1 Preattentive scanning versus attentive processing

Two mechanisms are thought to govern saccadic eye movements: a spatial system, which tells the eye where to go (the "where" system) and a temporal mechanism that tells the eye when to initiate a saccade (the "when" system). The "where" system may be the same spatial localization system that subserves perception [7, 14]. The "when" system is most likely located in the superior colliculus, consisting of both facilitatory and inhibitory neural circuits [8].

Findlay and Walker [3] present a model of a fixate/move system that triggers saccadic eye movements. The "move" command in the model is made at a relatively low level, governed by automatic processes as well as by higher-level voluntary control. A more complete picture of mechanisms involved is shown in figure 1 (see the color plate). One can easily see that these mechanisms are hierarchically organized and involve automatic as well as consciously controlled components.

Following a long tradition in cognitive psychology, Hoffman [6] uses terms of a two stage model of visual search. A preattentive stage of processing serves to locate objects in the visual world, to be examined by a second stage that is associated with attentive processing. This second stage is thought to operate only on a few objects at a time and forms a bottleneck in visual processing. Once there is attentive processing, the fixation duration is under voluntary control until further notice, which, in general, leads to longer fixations. This view is further supported by studies of Pomplun [9] and Velichkovsky et al. [18]. Accordingly, both fixation durations and dwell times seem to be under selective control of hierarchical mechanisms: during a visual search-and-comparison task, two phases could be distinguished on the basis of fixation durations. In the first (search) phase, fixation duration is mainly a function of spatial density. The second (comparison) phase, just prior to the solution, is considered to be characterized by hypothesis generation and testing. In this second phase, fixations are no longer controlled by spatial features of the task, but instead by the complexity of decision.

1.2 Distribution of fixation durations

During free viewing of a scene, fixation durations appear highly variable, ranging from less than 100 milliseconds to several seconds, and such changes even take place within consecutive fixations as has been shown by Buswell decades ago [2]. This is in marked contrast to saccadic reaction times, that are relatively stable, usually ranging from 150 to 250 ms. The distribution of fixation durations has a characteristic positive skew and a modal value that usually lies in the same range as the saccade reaction time process. Ulrich and Miller [10] described a number of simulation studies, showing that log-normal distributions may be the product of several conceptually different mechanisms of information processing. They distinguished between three possible causes of log-normality: exponentially transformed random normal variables, products of random normal vari-

ables, or a combination of both. Fixation distributions, being essentially non-normal, are thus likely to be produced by a number of processes, the duration of each may by itself be normally distributed.

How can we infer the preattentive scanning and attentive processing from looking at the distribution of fixation durations? Since modal fixation durations are stable in most settings, the modal value may at first sight not seem very informative. Yet considering that preattentive scanning prevails in most visual search tasks (e.g. during driving), and that preattentive processes are not subject to manipulations in the cognitive domain, it is highly likely that the distribution of fixations serving preattentive processes lies in the modal range and is not affected by cognitive factors. This supposition is further supported by the relation between saccade amplitude and fixation duration, as will be shown below.
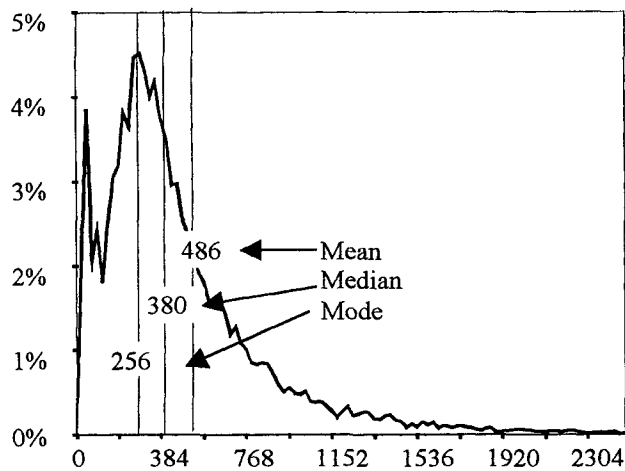

Figure 2: Typical frequency distribution of fixations as found during a simulated driving task (bin width 24 ms). A total of over 50.000 fixations pooled over all subjects and all conditions are taken from the experiment described below.

Let us consider a typical frequency distribution of fixation durations as found in our experiments (figure 2). The distribution shown is typical in the sense that it is log-normal, with the mode (256 ms) lying below the mean (486 ms). Furthermore, the distribution is not unimodal. Instead, a second, smaller mode appears around 60 ms. In our experiments, fixations in this area made up around 7% of all fixations. These extremely short fixations have previously been reported by Velichkovsky et al. [19]. They appear to indicate visuomotor readiness and are sensitive to changes in task conditions. The processes underlying these fixations are not yet fully understood, although its similarity to express-saccades (e.g. Fischer & Ramsberger [4]) makes the supposition attractive, that fixational disengagement plays a role. Following the line of argument used by Findlay and

Walker [3], this would either be because of a reduction of activity in the "fixate" centre (which is, within the express-saccade paradigm, usually the offset of the fixated target, and hence an unlikely candidate) or an increase in activity in the "move" centre. Usually, studies on fixation duration do not consider fixations of less than 100 ms duration, but since they make up a significant portion of the frequency distribution, it seems important not to ignore them.

Taking these considerations together, using the average fixation duration as a single parameter of attentional state seems not justified. Additional information about the distribution is not only necessary from a formal point of view, but also in order to refine the conclusions usually drawn on the basis of average fixation durations. The aim of this paper is to examine and compare quantitative and qualitative changes in fixation parameters due to task conditions, in particular with respect to preattentive search vs. attentive elaboration of critical events during driving.

2 METHODS

The results discussed here are derived from a study on eye movements and risk perception during a driving task performed on a PC-based driving simulator. The simulation consisted in the representation of a two-lane street with a quasi-random series of curves with varying radius. Simulation frame rate yielded approximately 40 frames per second, simulation data were stored at the same rate. Subjects were seated at a distance of 150 cm in front of a projection display subtending approximately 40° horizontally and 30° vertically. Eye movements were recorded with the Eyelink ™ head worn system with a temporal resolution of 250Hz.

Twenty four subjects (11 male, 13 female, with a mean age of 24.6 years, ranging from 20 to 34 years) took part in the study. The subjects were asked to drive at a constant speed, while keeping the car positioned in the middle of the right lane. Before starting the experiment, subjects were asked to complete a four minute test drive on the simulator in order to get accustomed to it. Three driving conditions were studied:

- baseline: driving at a constant speed, keeping position as closely as possible to the middle of the right lane, with no critical events;
- autopilot: the subject did not have to control lane position, but had to control speed, and brake upon critical events;
- standard: control both lane position and speed, brake upon critical events.

Three kinds of critical events were implemented: a lead car braking, a traffic light turning red and a ball rolling into the street. Non-critical events were: lead car driving, traffic light green, and a ball lying by the side of the road. Presentation of the critical events was timed in such a way that the location of the critical event would be reached 2500 ms from its appearance, assuming constant speed.

3 RESULTS

3.1 Driving condition

As already stated above, the same mean fixation duration may arise from quite different distributions. In our experiment, the mean fixation durations during the three driving conditions (see table 1) did not differ significantly from one another ($F_{2,24} = 0.673$, p=0.513). Post hoc analysis of the differences using a Scheffé Test did not reveal any significant differences between any of the three conditions.
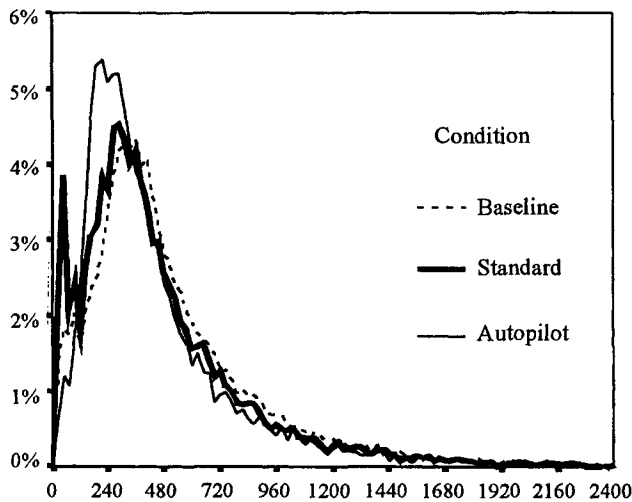
Table 1: Fixation duration (ms)

Condition	Mea	Mode
Baseline	543.	324
Autopilot	499.	216
Standard	508.	288

Figure 3: Distribution of fixation durations by driving condition. Only sections without further events are considered. Note the difference in modal duration between the baseline condition and the other two. Distributions are frequencies averaged by subject and condition, bin width is 24 ms.

Considering the means thus does not seem to reveal significant differences, whereas the difference between the distributions is marked (figure 3). In particular, there seems to be a marked differences between the autopilot and the other conditions in the range of modal values. A post hoc test of significance of the differential contribution of this range of fixation durations was performed on a subset termed "modal" fixations of 216 ± 48 ms. Multivariate repeated measures analysis of variance revealed significant differences between the driving conditions ($F_{2,23}$=74.09, p<0.001). Differences between autopilot and the other two conditions are significant at 5% level after Bonferroni correction. It seems therefore that these "modal" fixations are prolonged when participating in the sensomotor activity of lane-keeping, whereas merely observing (autopilot condi-

tion) does not lead to a prolongation of preattentive fixations.

It should be noted that both averages and modal values in this experiment were higher than corresponding parameters from studies of static pictures. Two possible causes for this difference come to mind. One obvious possibility is, that smooth pursuit ("dynamic fixations") has been activated in view of motion information. Another option is, that the deictic strategies (see e.g. Ballard et al. [1]; Velichkovsky [15]) of using the display as an external memory is not applicable in dynamic situations, which raises the need to process more data.

3.2 Saccade amplitude and fixation duration

In the previous section we considered evidence for the assumption that modal fixations are indicative of the preattentive state by showing differential effects of task conditions on classes of fixation durations. More evidence for the assumption is given by the fact, that modal fixations are usually followed by larger saccades than any other category.
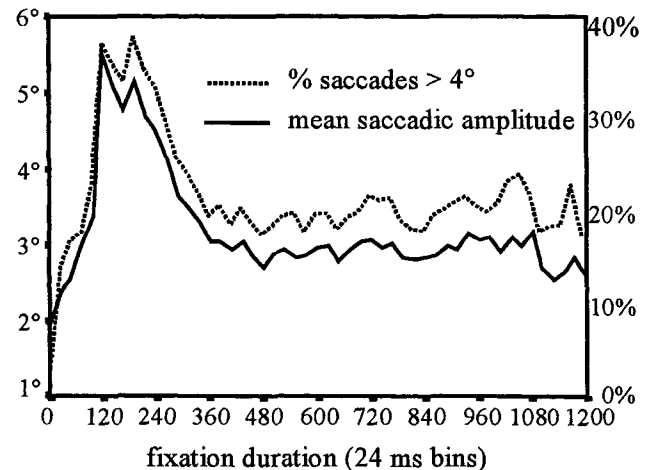

Figure 4: Relation between saccade amplitude and fixation duration. The saccades shown are those following the fixation.

Figure 4 clearly shows that the largest saccades are those being generated during fixations of modal duration. Furthermore, the percentage of saccades larger than 4° is also largest in the modal range. Combining these results with the fact that modal fixations were the most prominent in the autopilot condition, it is not surprising, that saccadic amplitude (F2,23=12.10, p<0.001) and the percentage of saccades over 4° (F2,23=16.65, p<0.001) also significantly differ between driving modes. Scheffé tests revealed significant differences between autopilot and the other two conditions at the 5% level for both the mean amplitude and the percentage of saccades over 4°.

Obviously, the modal fixations accompany saccade amplitudes that are clearly above average. If fixations serving preattentive search are indeed those that precede saccades of above average amplitudes, a selection of those fixations should thus yield a distribution that is centred around the mode, whereas the remainder of the fixations should have a modal value that lies well beyond that. Figure 5 shows three sets of fixations, grouped according to their subsequent saccade amplitude. One set consists of fixations preceding extremely small saccades (< 0.25°). Remarkably, more than 60% of these fixations fall within the express-fixation range. Express-fixation distribution thus generally seem to precede extremely short saccades. The second set is built by fixations preceding a saccade of 4° amplitude or larger. Their distribution clearly peaks in the modal range (240 ms), and appears to be rather more symmetrical than the original distribution. The third group of fixations, consisting of the remainder, has a modal value lying around 360 ms and is also considerably less skewed than the distribution they are taken from. Though highly speculative, it seems that the distribution of fixations serving preattentive scanning can be dissociated from those serving attentive elaboration.
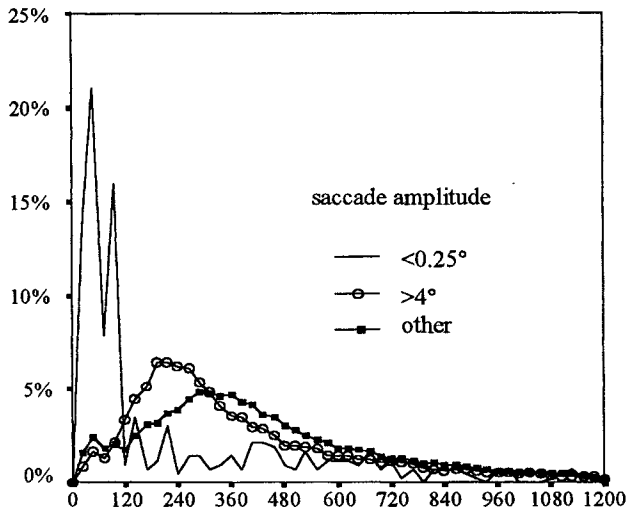

Figure 6: Distribution of fixation duration by levels of critical event. Frequencies averaged by subject and event, bin width 24 ms.

The data presented in figure 6 represent the averages over all critical events (lead car braking, traffic light red and ball rolling), averaged over the whole duration of the critical event.

Since fixation durations may change instantaneously from one fixation to the next, the duration of the fixation that actually "detects" the critical event is deeply embedded in this way of analyzing fixation durations. We therefore selected the time of braking as a starting point, and took a closer look at fixations occurring around this time.

Figure 7a shows the average fixation duration plotted over the fixation number relative to the time of braking.

Figure 5: Fixation duration distributions grouped by subsequent saccade amplitude.

3.3 Critical events

What happens to fixation durations when a critical event occurs? We selected periods during which critical situations occurred with those without event. Two-way analysis of variance again revealed no significant differences in average fixation duration between driving mode, but a significant difference (F2,23=3.90, p=0.05) between situations with and without critical events. Figure 4 shows the differences between the two distributions. Note that during a critical event, the fixations in the (presumably attentive) range of 300 to 500 ms are more frequently represented, whereas those in the modal range are less represented.
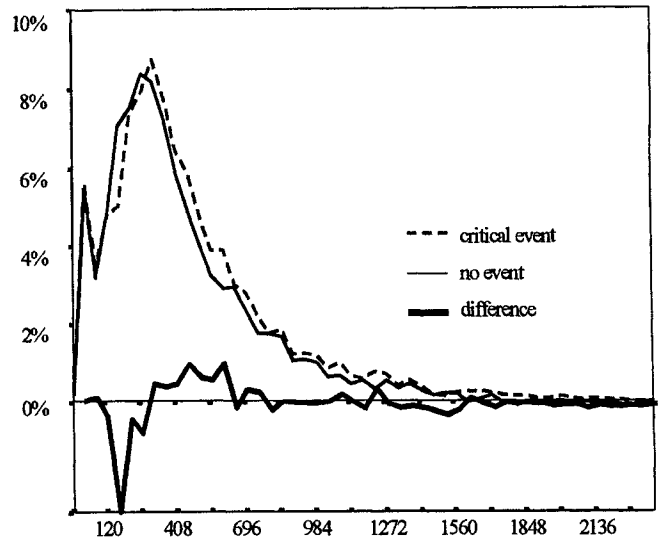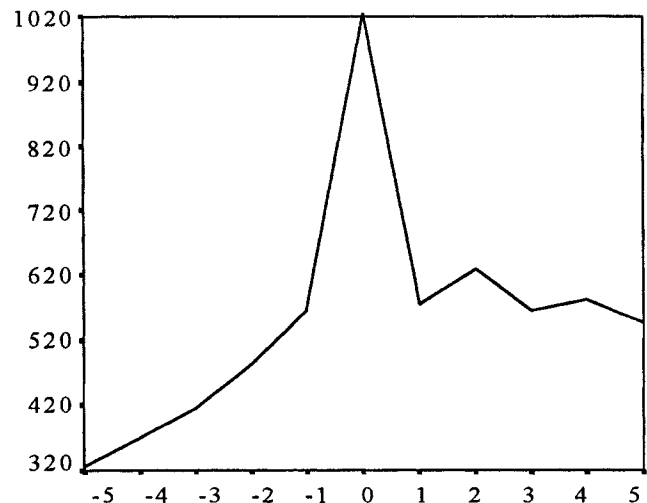

Figure 7a: Mean fixation durations around the time of braking. Durations are given in milliseconds. Zero corresponds to braking time.
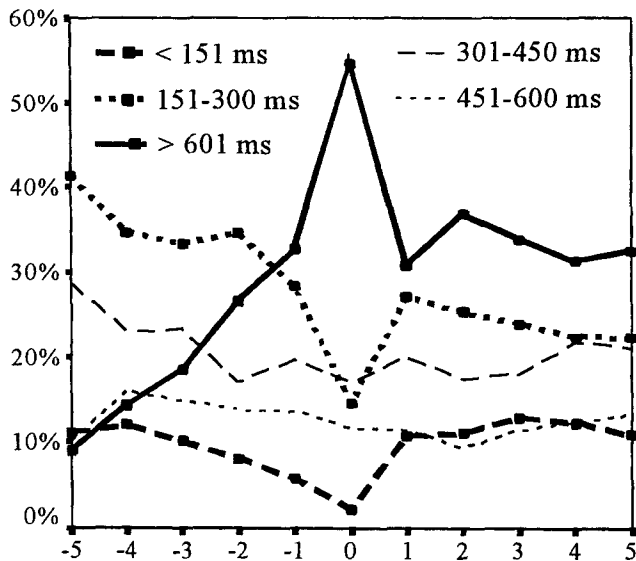
Figure 7b: Frequencies of fixation durations around the time of braking upon a critical event. Categories totalling 100%.

Figure 7b furthermore reveals clearly that some categories of fixation durations are more affected by the event than others: there is a considerable reduction of short fixations (categories 0-150 ms and 151-300 ms) and a marked increase of longer fixations - 600 ms and more. The two inbetween categories of 301-450 ms and 451-600 ms, on the other hand, seem not to be affected by the critical events

4 DISCUSSION

The results presented in this study show that analyzing fixation durations has implications both for theoretical and applied research. The distinction between preattentive scanning and attentive elaboration in visual search seems to be supported by the data, though these empirical results are even richer than such a dichotomic distinction (on the perspective of a multilevel architecture in uinvestigation of the cognitive aspects of eye movements - see [16]). The two-stage search process can be recognized by inspection of distribution data. A statistical analysis of multimodal distributions has not been performed, however. One possible way to test for the existence of basis functions underlying the distribution has been proposed by Gezeck and Timmer [5]. Their method seems particularly useful to study serially dependent processes.

The fact that these data result from an applied setting rather than from synthetic laboratory tasks emphasizes their importance for applied purposes. The results show that an estimate of level of attentional processing on the basis of eye movement analysis is within reach. We realize, however, that some of the analyses done are post-hoc, and need further support by hypothesis-testing rather than explorative studies. In particular, the difference in expression of

modal values between autopilot and driving modes in which maintaining lane-position must be controlled seems to indicate that the act of steering is an attention demanding rather than an automated process. On a preliminary basis, we propose the idea, that fixations subserving preattentive scanning are clearly dissociable from those serving attentive elaboration on the basis of their duration as well as the subsequent saccade amplitude. Of practical importance is that this information can be obtained without taking into account data about exact spatial location of eye movements.

The analysis of fixational behaviour during critical events, finally, showed that a shift in the processing level from preattentive to attentive is recognizable on a very short, phasic time scale (see also Unema & Velichkovsky, in press). Particularly prominent are the sudden increase in fixation duration upon detection of a critical event and the corresponding increase in occurrence of long fixations on the cost of fixations with modal durations, i.e. durations of about 200 ms.

REFERENCES

[1] D.H. Ballard, M.M. Hayhoe, P.K. Pook, and R.P.N. Rao. Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, 20: 723-767, 1997.

[2] G.T. Buswell. *How People Look At Pictures.* Chicago, University of Chicago Press. 1935.

[3] J.M. Findlay and R. Walker. A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, 22: 348-362, 1999.

[4] B. Fischer and E. Ramsberger. Human express saccades: extremely short reaction times of goal directed eye movements. *Experimental Brain Research*, 57, 191-195, 1984.

[5] S. Gezeck, and J. Timmer. Detecting multimodality in saccadic reaction time distributions in gap and overlap tasks. *Biological Cybernetics*, 87: 293-305, 1998.

[6] J. E Hoffman. Stage of processing in visual search and attention. In B.H.Challis and B.M. Velichkovsky (Eds.), *Stratification in Cognition and Consciousness*, pp. 43-71, Amsterdam/Philadelphia: John Benjamins. 1999.

[7] E. Kowler. Eye movements. In S. Kosslyn and D.N. Osherson (Eds.), *Visual Cognition*, pp. 214-265, Camgbridge, MA: MIT Press.1995.

[8] D.P. Munoz and R.H. Wurtz. Fixation cells in monkey superior colliculus I: Characteristics of cell discharge. *Journal of Neurophysiology*, 70: 559-575. 1993.

[9] M. Pomplun. *Analysis and Models of Comparative Visual Search.* Aachen: Cuvilier. 1998.

[10] R. Ulrich and J. Miller. Information processing models generating lognormally distributed reaction times. *Journal of Mathematical Psychology,* 37: 513-525. 1993.

[11] P. Unema. *Eye Movements and Mental Effort.* Aachen: Shaker. 1995.

[12] P. Unema and B.M. Velichkovsky. Processing Stages as Revealed by Dynamics of Visual Fixations: Distractor versus Relevance Effects. Paper presented to the 41st Annual Meeting of Psychonomic Society. 2000.

[13] T. Van-Zandt and R. Ratcliff. Statistical mimicking of reaction time data: Single-process models, parameter variability, and mixtures. *Psychonomic Bulletin and Review,* 2: 20-54, 1995.

[14] B.M. Velichkovsky. The vertical dimension of mental functioning. *Psychological Research,* 52: 282-289. 1990.

[15] B.M. Velichkovsky. Communicating attention: Gaze position transfer in cooperative problem solving. *Pragmatics and Cognition,* 3(2): 199-222. 1995.

[16] B.M. Velichkovsky. From levels of processing to stratification of cognition. In B.H.Challis & B.M.Velichkovsky (Eds.), *Stratification in Cognition and Consciousness.* Amsterdam/Philadelphia: John Benjamins. 1999.

[17] B.M. Velichkovsky and J.P. Hansen. New technological windows into mind: There is more in eyes and brains for human-computer interaction.(Keynote address) In CHI-96: *Human Factors in Computing Systems.* New York: ACM Press. 1996.

[18] B.M. Velichkovsky, B.H. Challis and M. Pomplun. Arbeitsgedächtnis und Arbeit mit dem Gedächtnis: Visuell-räumliche und weitere Komponenten der Verarbeitung [Working memory and work with memory: Visual-spatial and further components of processing]. *Zeitschrift für experimentelle Psychologie,* 42: 672-701. 1995.

[19] B.M Velichkovsky, A. Sprenger and P. Unema, Towards gaze-mediated interaction: Collecting solutions of the "Midas Touch Problem" In S.Howard, J.Hammond & G. Lindgaard (Eds.), *Human-Computer Interaction: INTERACT'97* (Sydney, July 14-19th), Chapman & Hall, London. 1997.
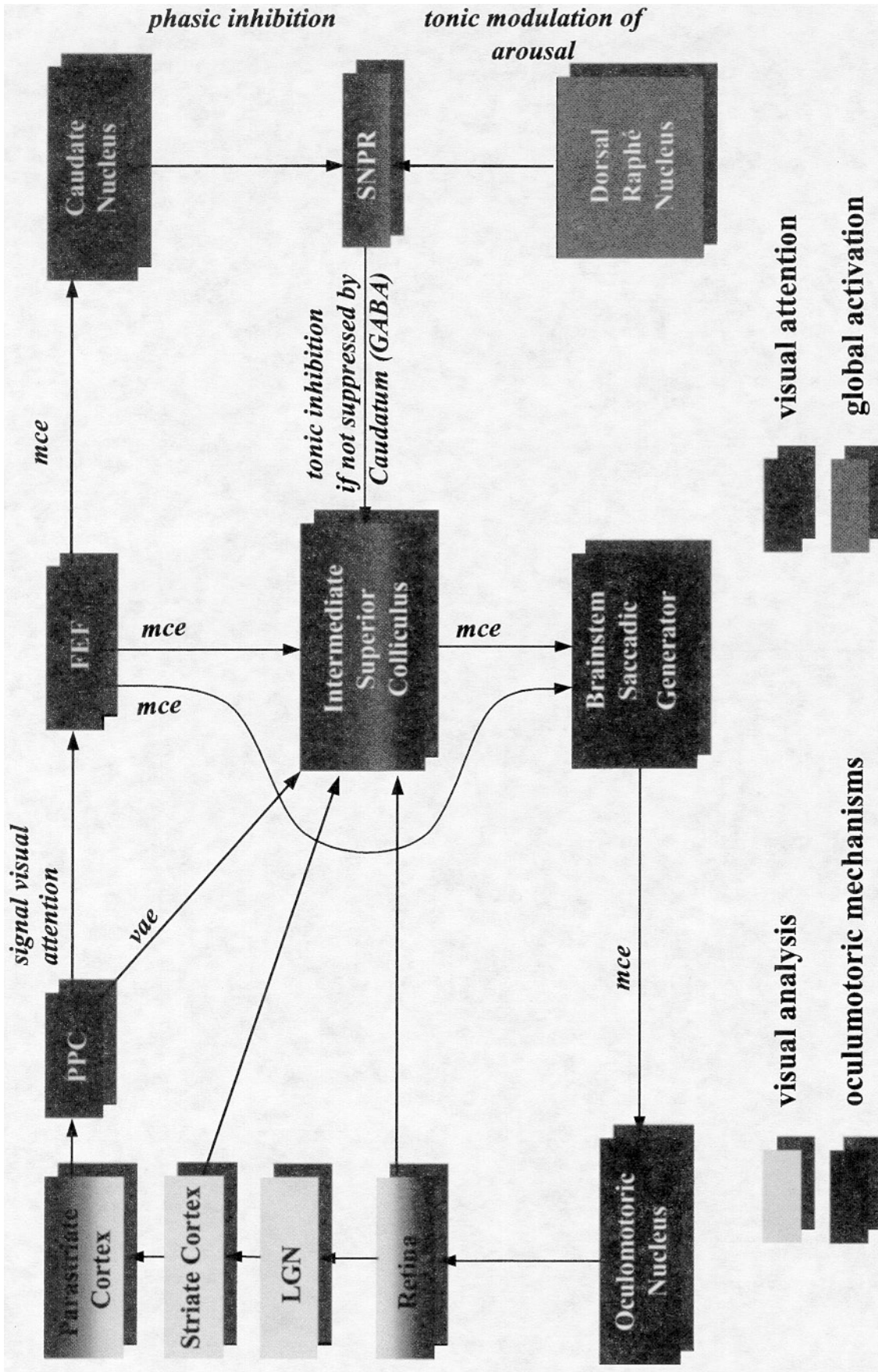
Figure 1 Neurobiological mechanisms of eye movement control: PPC, posterior parietal cortex; FEF, frontal eye field; LGN, lateral geniculate nucleus; SNPR, substantia nigra pars reticulata; vae, visual attentional excitatory; mce, motor command excitatory,