



UPPSALA  
UNIVERSITET

*Digital Comprehensive Summaries of Uppsala Dissertations  
from the Faculty of Social Sciences 9*

# Keeping Eye and Mind on the Road

TRENT VICTOR



ACTA  
UNIVERSITATIS  
UPSALIENSIS  
UPPSALA  
2005

ISSN 1652-9030  
ISBN 91-554-6425-4  
urn:nbn:se:uu:diva-6241

Dissertation presented at Uppsala University to be publicly examined in Sal X, Universitetshuset, Uppsala, Friday, January 20, 2006 at 13:00 for the degree of Doctor of Philosophy. The examination will be conducted in English.

**Abstract**

Victor, T. 2005. Keeping Eye and Mind on the Road. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 9. 83 pp. Uppsala. ISBN 91-554-6425-4.

This thesis is devoted to understanding and counteracting the primary contributing factor in traffic crashes: inattention. Foremost, it demonstrates the fundamental importance of proactive gaze in the road centre area for action guidance in driving. Inattention is explained with regard to two visual functions (vision-for-action and vision-for-identification), three forms of attentional selection (action-driven-, stimulus-driven-, and goal-directed attention), and two forms of prediction influences (extrapolation-based- and decision-based prediction influences). In Study I an automated eye-movement analysis method was developed for a purpose-built eye-tracking sensor, and was successfully validated. This analysis method was further developed, and several new measures of gaze concentration to the road centre area were created. Study II demonstrated that a sharp decrease in the amount of road centre viewing time is accompanied by a dramatic spatial concentration towards the road centre area in returning gaze during visual tasks. During cognitive tasks, a spatial gaze concentration to road centre is also evident; however contrary to visual tasks, road centre viewing time is increased because the eyes are not directed towards an object within the vehicle. Study III found that gaze concentration measures are highly sensitive to driving task demands as well as to visual and auditory in-vehicle tasks. Gaze concentration to the road centre area was found as driving task complexity increased, as shown in differences between rural curved- and straight sections, between rural and motorway road types, and between simulator and field motorways. Further, when task duration was held constant and the in-vehicle visual task became more difficult, drivers looked less at the road centre area ahead, and looked at the display more often, for longer periods, and for more varied durations. In closing, it is shown how this knowledge can be applied to create in-vehicle attention support functions that counteract the effects of inattention.

*Keywords:* Vision, Eye-movements, Driving, Attention, Distraction, Ventral stream, Dorsal stream, Road safety, Traffic crashes, Evaluation methodology

*Trent Victor, Department of Psychology, Box 1225, Uppsala University, SE-75142 Uppsala, Sweden*

© Trent Victor 2005

ISSN 1652-9030

ISBN 91-554-6425-4

urn:nbn:se:uu:diva-6241 (<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-6241>)

*To Gunilla, Ossian, Rasmus, and Stella*



## List of Papers

This thesis is based on the following papers, which will be referred to in the text by their Roman numerals:

- I Victor, T., Blomberg, O., & Zelinsky, A. (in press). Automating driver visual behavior measurement. *Vision in Vehicles 9*.
- II Victor, T. W, Johansson, E. (2005). Gaze concentration in visual and cognitive tasks: Using eye movements to measure driving information loss. Manuscript submitted for publication. Uppsala University.
- III Victor, T. W., Harbluk, J. L., & Engström, J. A. (2005) Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F: Psychology and Behaviour*. 8(2), 167-190. DOI: 10.1016/j.trf.2005.04.014



# Contents

Introduction.....	9
The problem of inattention in driving .....	9
Active Vision.....	10
The strong illusion of a stable, detailed image of the external world ..	10
Multiple concurrent processes, specific functions, blended processing steps .....	12
Two-stream vision as a guiding principle – vision-for-action and vision-for-identification.....	14
Vision, attention, and eye movements .....	19
Eye movements in driving.....	20
Attention.....	25
Attention is selection .....	26
Stimulus-driven attention and goal-directed attention.....	27
Action-driven attention.....	28
Prediction – Expectation and Extrapolation.....	30
Empirical studies.....	32
General aims.....	32
Automated and validated visual behavior measurement (Study I).....	33
The initial eye-movement analysis algorithm.....	36
Results .....	36
Discussion.....	37
Gaze concentration and other eye movement effects of naturalistic tasks (Study II) .....	38
Design.....	39
Results .....	40
Discussion.....	46
Eye movement effects of different task types, task difficulty, and driving complexity (Study III).....	51
Design.....	51
Results .....	52
Discussion.....	56
Summary.....	59
General Discussion .....	60
Summary of main empirical contributions.....	60

Performance dissociations in path control and reaction time measures are associated with eye movements .....	61
Main contributing factors to inattention in crashes .....	63
Attention Support – How to keep eye and mind on the road in practice .....	68
Distraction alerts .....	68
Inattention sensitive driving support functions .....	71
Peripheral information displays .....	71
Path keeping displays .....	71
Workload Management .....	73
Acknowledgements .....	74
References .....	75

# Introduction

The main focus of this thesis is to understand the problem of inattention in driving. Several sections are thus devoted to understanding recent key developments within vision, eye movements, attention, and prediction as a background to the empirical work. A second aim is to apply this knowledge to the results of the empirical work to explain how visual performance is affected by various activities while driving. In closing, it is shown how this knowledge can be applied to in-vehicle products to improve safety.

## The problem of inattention in driving

Loss of forward roadway vision, even for a split second, can lead to disastrous consequences in driving. In a recent, very significant study of large-scale naturalistic driving data, 'inattention to the forward roadway' was found to be the primary contributing factor in crashes (Neale, Dingus, Klauer, Sudweeks, and Goodman, 2005). 78 percent of all crashes, and 65 percent of all near-crashes involved driver inattention just prior (within 3 seconds) to the onset of a conflict. 93 percent of all rear-end-striking crashes involved 'inattention to the forward roadway'. The study produced 42,000 hours of video- and driving-performance recordings over 12 months from more than 100 drivers. Within this data, the causal and contributing factors behind 15 police-reported crashes, 67 non-police reported crashes, 761 near-crashes, and 8,295 incidents were recorded. Driver inattention encompassed four categories: 'secondary task engagement', 'driving-related inattention to the forward roadway', 'non-specific eye glance away from the forward roadway', and 'fatigue'. The first three categories involve looking away from the forward roadway, and the last category involves loss of forward roadway vision from eyelid closure. Thus, glances away from the forward roadway were found to contribute to a much greater percentage of events than had been previously thought. These results also lend strong support to the model, devised by Wierwille and Tijerina (1998), which almost perfectly predicted real-life crash rates from off-road-glance behavior and device use.

Further, expectancy was found to be an important simultaneous factor, as drivers had difficulty responding appropriately when other vehicles perform unexpected or unanticipated maneuvers and when expectancies about the flow of traffic are violated, such as sudden stops or lane changes. Green

(2000), Evans (2004), and Rumar (1990) also point to expectation as a crucial factor influencing reaction times.

The 100-car study statistics greatly increase the significance of inattention in crashes. For example, the previous estimate of inattention in crashes by the United States National Highway Traffic Safety Administration (NHTSA) was at 25% to 30% of police-reported traffic crashes, or 1.2 million crashes per year (Shelton, 2001). In another widely-cited study (Treat, et al., 1977), perception and comprehension problems, categorized as “recognition errors”, predominated in the human direct causes that led to the conclusion that human factors are involved in 92.6% of accidents. The recognition errors categories were improper lookout (23% of accidents), inattention (15%), and internal distraction (9%). Major types of improper lookout included drivers changing lanes, passing, or pulling out without looking carefully enough for oncoming traffic. Major examples of inattention were a delay in detecting that traffic ahead was either stopped or decelerating, and a failure to observe critical road signs and signals.

In sum, the two key causal factors are combined in the simultaneous occurrence of a) loss of forward roadway vision, and b) an unexpected event. In the next sections we will see that we lose forward roadway vision in a variety of ways – when we move our eyes, when we blink, and while we are looking away, but also as a result of cognitive factors. An understanding of the problem of inattention in crashes and an understanding how to develop effective crash countermeasures arises out of knowledge about vision, attention, and prediction.

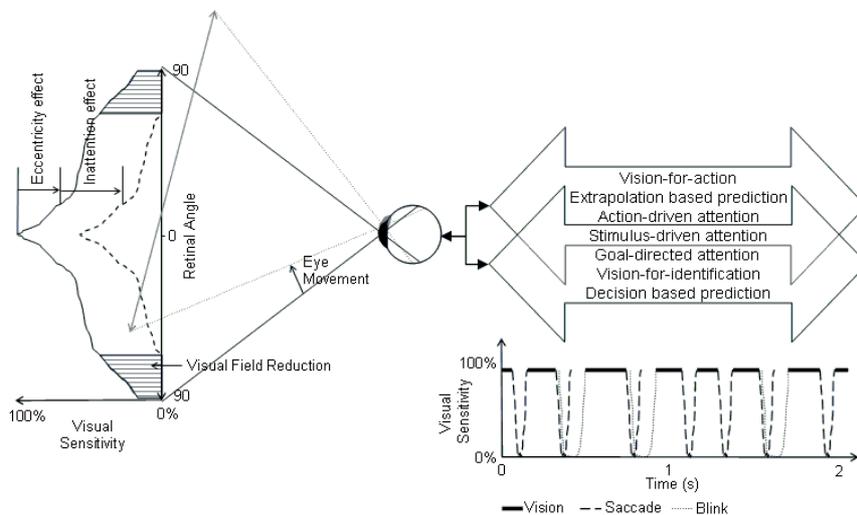
## Active Vision

The field of vision research has taken great strides with the onset of modern research technology, however the relevance of many of these results are relatively unknown to traffic researchers and human factors specialists developing in-vehicle information and communication systems and advanced driver assistance systems. This section aims at laying a modern foundation for understanding vision in driving.

### The strong illusion of a stable, detailed image of the external world

The illusion of a stable, full detail, pictorial view of the external world is very powerful and gives us a strong subjective impression of seeing everything when in fact we do not (Findlay & Gilchrist, 2003; Kalat, 2004; O’Regan & Noë, 2001). This illusion is created mainly because of our ability to direct our eyes so quickly and effortlessly to any desired location, thereby

giving the impression that we see detail across the full visual field despite a dramatic reduction in the performance of many visual functions towards visual field periphery (Findlay & Gilchrist, 2003; Leigh & Zee, 1999), see Figure 1. The eyes are not cameras that deliver a uniformly detailed picture image. This reduction in performance towards the periphery is caused by a decreasing density of receptors in the retina, but also because central visual regions receive an increasingly higher proportion of cortical processing in higher cortical regions (Findlay & Gilchrist, 2003). The fovea is not simply an area of high acuity, but rather the location at which visual processing is centered. Therefore, moving the eyes necessarily gives more processing capacity to the fixated region. This sensory limitation is easily mistaken as a central limitation, for example by Averbach and Coriell (1961) as described in van der Heijden (1996). An active vision approach emphasizes a dynamic view of the process of seeing, with a particular emphasis on visual attention, eye movements, and the importance of retinal inhomogeneity in vision (Findlay & Gilchrist, 2003).



*Figure 1.* Conceptual model of key factors influencing inattention in driving. Visual sensitivity is dependent on the angle from the fovea whereby a dramatic reduction in the performance of many visual functions occurs toward the periphery (eccentricity effect). Sensitivity across the retina is also affected by inhibitory effects of inattention (inattention effect), also resulting in visual field size reduction. The visual signal across time is only available between saccades and blinks and is used by two main, somewhat overlapping functional purposes – a) continuous on-line control of action processes (vision-for-action, extrapolation based prediction, action-driven attention, and stimulus-driven attention) and, b) discrete higher cognitive processes (vision-for-identification, decision based prediction, goal-directed attention, and stimulus-driven attention).

The stable, detailed image illusion is also maintained despite the fact that our eyes are like shutters. Vision is only available to us *between* saccades and blinks. Saccades, our fast eye movements, occur three to four times a second or more than 100 000 times a day, and blinks occur once or twice a second (Leigh and Zee, 1999). Every time the eyes make a saccade, a mechanism called saccadic suppression inhibits vision so that we cannot see the motion or blurred image caused by moving the eyes (Leigh and Zee, 1999). Saccadic suppression is the reason why you cannot see your own eyes move in the mirror (try it!). The larger the saccade is, the longer the period of vision loss. So, when a driver looks back and forth between the road and a low placed display, the amount of saccadic-suppression-time is larger than with a highly placed display. Saccadic suppression and blinks also play an important role in change blindness as we will see.

In addition to these sensory limitations to vision, there are limitations in internal representations of the outside world. Although the external world seems to be an integrated whole, different aspects such as shape, color, speed and direction of movement are being analyzed much more independently in different parts of the brain than has been imagined (Kalat, 2004). Perhaps the strongest incompatible results to the ‘picture in the head account’ are given by the Change Blindness paradigm, which show that visual representations are limited and that we need attention to explicitly “see” (e.g. Rensink, 2002a). It is only when we focus our attention on a part of the visual field that our experience becomes fully informed about what is actually out there (Rensink, 2002a). The visual world is believed to act as an external memory, implying that complex internal representations are unnecessary because information can be acquired on demand rather than by relying on an internal model, representations are transient and geared to the immediate task (O’Regan and Noë, 2001; Rensink, 2000). This is in strong contrast to the traditional approach to vision, whereby the brain was thought to reconstruct a general-purpose representation of the information in the scene (e.g. Marr, 1982).

Just as most drivers are completely unaware these limitations, many theories overlooked or ignored them. The legacy of “the mental picture in the head” account still lingers in many dark corners, including theories of driving.

### Multiple concurrent processes, specific functions, blended processing steps

Although it is tempting to describe behavior in three steps – perception, cognition, action – the brain does not handle the process in such discrete steps (Kalat, 2004; Goodale, 1998). Modern cognitive neuroscience supports a more “specific functions, blended processing steps” view on behavior. Be-

havior emerges from an enormous number of partly independent, partly interdependent processes occurring simultaneously throughout the nervous system (Kalat, 2004). Complexity frequently takes the form of a hierarchy, so it is not surprising that the multitude of processes is dealt with using hierarchical design as a general principle of cortical architectonics (Koch, 2004). The well-known Felleman and Van Essen (1991) chart shows that the visual system is well ordered to a highly stratified hierarchy, interconnected by a few hundred linkages, most of which are reciprocal.

To illustrate independence of visuomotor modules, consider the following example. Ingle (1973) rewired a visuomotor control module, causing a frog to stick its tongue out in the opposite direction when trying to catch prey, even though it maintained the ability to maneuver around visual obstacles blocking its path. What did the frog see? A mirror image when snapping at the prey and a correct image when avoiding barriers? If you accept that there are different visuomotor modules in the brain the puzzle disappears. In fact, Milner and Goodale (1998) and Goodale and Milner (2004) argue that each part of the animal's behavioral repertoire has its own separate visual control system with independent pathways from visual receptors to the motor nuclei. They assert that vision evolved originally to provide distal sensory control of movements, and that only later did representational systems evolve to permit the brain to model the world, identify object and events, and to attach meaning, causality, and significance to them (Milner & Goodale, 1998). The representational systems are thus linked to cognitive systems subserving memory, semantics, planning, and communication (Milner & Goodale, 1998).

The perceptual representations can be used for many different purposes, be shaped by memories, emotions, and expectations. They allow us to choose a goal, plan ahead, and decide on a course of action. But on the other hand they do not have any direct contact with the motor system. The role of the perceptual representational systems is not in the execution of actions but rather in helping the animal arrive at a decision (Goodale & Milner, 2004). The representations used by the perceptual representational systems are different than those used by the real-time visuomotor systems.

So, although you may think that you see a single integrated representation of the visual scene, your brain is actually analyzing different aspects of the scene separately using different quasi-independent modules, both in terms of the networks that mediate object classification, recognition, and identification, and in terms of modules for specific behaviors, such as eye movements, posture adjustment, hand gripping or pointing, foot placement, and so on. Importantly, this modularity is based more on the nature of the actions guided by vision (e.g. reaching, saccades, whole-body locomotion) than on visual features (Goodale and Milner, 2004). Each action element may be individually guided by different visual information. This position clearly refutes the common assumption that behavior is controlled using a single general-purpose representation of the visual world.

The representations used by the different visuomotor systems can be graded in strength (Munakata, 2001; Jonsson & von Hofsten, 2003). Graded representation theory states that some actions will suffer more from poor strength of a representation than others. Some actions may only require weaker representations, while other actions may require stronger representations to function properly. For example, reaching requires more precise representations than eye/head tracking. This is important because representations lose strength and precision with time, for example when looking or attending away from an object. This implies that some visuomotor systems used while driving may be more sensitive to a loss of representation from inattention than others.

### Two-stream vision as a guiding principle – vision-for-action and vision-for-identification

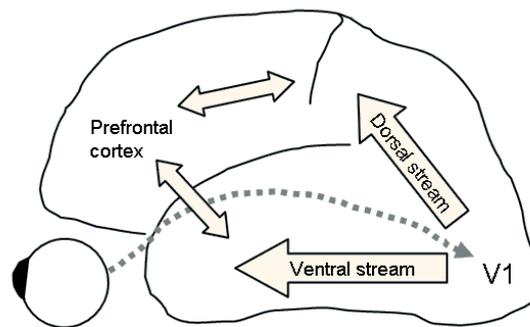
Driving is a complicated activity, yet amazingly, we sometimes do it automatically without thinking about it. We can drive all the way home by mistake when we wanted to go to the supermarket. Likewise, we can become engaged in thought or a mobile phone to the point where we stop being aware of the fact we are also driving. For example, one of my subjects was shocked to learn that he inadvertently passed the truck he was instructed to follow while performing an arithmetic task. Try introspecting next time you drive, you will find yourself steering and looking at the road without any knowledge of which mechanism caused the steering corrections to the appropriate amount.

The point made previously is that there is not just one way of seeing. The visual information you explicitly experience is not the same as the visual information that guides your actions. In fact, much of what vision does for us lies outside our visual experience. Although we become aware of the actions that the visuomotor systems control after the fact, we have no direct experience of the visual information they actually use to do so. Information enters the nervous system and influences action even when it doesn't gain conscious attention.

A recent cornerstone in visual neuroscience has far-reaching implications for understanding driving. As discovered by Ungerleider and Mishkin (1982), vision is based on two broadly separate, but interactive, cortical routes of processing, called the dorsal and ventral visual streams. The streams use different representations and transform information for different functional purposes. The dorsal stream underlies the control of actions, and the ventral stream is responsible for identifying and representing enduring characteristics of objects (Goodale & Westwood, 2004; Milner & Goodale, 1995), see Figure 2. This division of labor is one of the most influential concepts in understanding vision.

Evidence for Milner and Goodale's interpretation of the two streams (Milner & Goodale, 1995) builds on evidence from neuroanatomy, neurophysiology, neuropsychology, eye movement studies, and psychophysics, with special emphasis on the dissociations between identification and the visual control of action that may be observed, in both monkeys and humans, after damage to different visual areas of the brain (Goodale & Westwood, 2004). Recently, the functional specialization of the two streams has clearly been demonstrated in intact humans (Shmuelof & Zohary, 2005).

This division of labor is believed to have evolved because of the different processing demands imposed by the two purposes of vision (Goodale & Milner, 2004; Koch, 2004). The two streams represent the way our nervous system solves the task of interacting with the visual world. The mere fact that the brain divides the labor in this way means that the two streams solve independently important and complementary aspects of behavior. The two streams thesis provides an appropriate model and framework for understanding the relationship between action control mechanisms and more cognitive mechanisms that are used for driving.



*Figure 2.* The dorsal, vision-for-action and ventral, vision-for-identification streams in relation to various regions of the brain. The flow of visual information splits in V1 into two streams that diverge and flow toward either the interior temporal cortex (ventral stream) or the posterior parietal cortex (dorsal stream). From there they project to different parts of the prefrontal cortex, where they reconverge. While the ventral stream handles form and object recognition, the dorsal stream carries spatial information for locating targets and executing motor actions. Modified from Koch (2004).

### **Vision-for-action – the dorsal stream**

Action control requires fast processing and spatially accurate responses. The nervous system has solved this problem by having the dorsal stream transform visual information directly, through non-conscious vision-action links. Vision-for-action delivers accurate metrical information regarding size, location, and motion in egocentric, viewpoint-dependent, absolute coordinates to

the visuomotor systems (Milner & Goodale, 1995; Goodale & Milner, 2004). For example, eye movements use retinocentric coordinates, and reaching uses head or body-centric coordinates. Thus, visuomotor systems are given access to the true real-world position of a target relative to the body.

Computations are short-lived and do not need to access working memory. The vision-for-action, visuomotor control systems work in real time. When a delay is imposed, our conscious visual experience intrudes. A large amount of converging evidence shows that sensorimotor representation can only be expressed within a short delay following stimulus presentation (see Rosetti & Pisella, 2002, for a review). The general finding in normals and patients in manipulation tasks is that a 2 second delay strongly affects motor output performance. Rosetti and Pisella suggest a continuum between a pure sensorimotor system and a pure cognitive identification system, and that the crucial parameter of this gradient is the amount of on-line processing participating in a given task.

Similar to Goodale and colleagues, Koch and Crick (2001) and Koch (2004) assert that the sensorimotor modules in vision-for-action stream, which they call “zombie agents”, function in the absence of consciousness. We become conscious of the action of a sensorimotor module if we attend to it, but usually only after the fact, through delayed internal or external feedback (Koch, 2004). Castiello, Paulignan, and Jeannerod (1991) estimated the delay between the onset of a rapid manual response and subjective awareness to be 250 ms. This means, for example, that a sprinter is already out of the blocks before consciously hearing the starting gun.

In sum, vision-for-action uses a set of dedicated, fast, metrically precise sensorimotor modules to carry out just-in-time calculations that convert visual information directly into action (Milner & Goodale, 1995; Goodale & Milner, 2004; Goodale & Westwood, 2004).

### **Vision-for-identification – the ventral stream**

Our conscious visual experience is a product of the ventral, vision-for-identification stream, as indicated by converging evidence from lesion evidence, fMRI evidence, and single neuron recordings (Goodale & Westwood, 2004; Koch, 2004; Shmuelof & Zohary, 2005). Vision-for-identification is a slower, general purpose visual stream that delivers a rich and detailed representation of the world. It mediates form, orientation, color, face and object recognition, classification, identification, and goals for the vision-for-action stream (Milner & Goodale, 1995; Goodale & Milner, 2004). Importantly, it throws away the detailed metrics of the scene with respect to the observer. Because these representations are used as a foundation for thinking about the past, present and future visual worlds, they are in a form that is viewpoint-independent. Vision-for-identification uses an object-centered, scene-based spatial map, with relative metrics (Milner & Goodale, 1995; Goodale &

Milner, 2004). It allows for indefinitely long processing time and continues to work after long delays because of its access to memory.

Vision-for-identification transformations operate on highly processed, learned visual information that take into account previous knowledge and experience in long term memory to form representations of the enduring characteristics of objects. Thus, our conscious visual experience is derived as much from memory as from visual input. As the complexity of actions increases and goals extend in space and time, we depend less on automatic visuomotor systems and more on interactions with long-term memory, decision making, and other cognitive processes (Koch, 2004).

### **How the two visual streams work together**

The two streams evolved together and play complementary roles in the control of behavior. There is good evidence from brain anatomy that the two streams are interconnected (Goodale & Milner, 2004). In the normal course of everyday events, the functioning of vision-for-action and vision-for-identification is tightly interwoven. In general, vision-for-action systems operate very rapidly without conscious awareness, but their goals are primarily set by vision-for-identification (Goodale & Milner, 2004; Milner & Goodale, 1995; Koch, 2004). Vision-for-identification helps us plan and identify the goals for action, the class of action to perform, and even helps tune action parameters based on prior knowledge, for example by deciding the scaling of initial lifting forces from prior experience (Milner & Goodale, 1995; Goodale & Milner, 2004). When a particular goal object has been flagged by vision-for-identification, dedicated visuomotor networks in vision-for-action, in conjunction with output systems elsewhere in the brain, perform the desired motor act (Goodale & Milner, 2004). So vision-for-identification is involved making contributions to action not just at the planning stage, but right down to the programming of force.

The relationship has been likened by Goodale and Milner (2004) to the way a human operator operates a semi-autonomous robot in teleassistance, for example when controlling a vehicle on Mars. In teleassistance the human operator has a job of identifying a goal and specifying an action and the robot vehicle finds its way to the goal.

An excellent example illustrating both cooperation and functional separation comes from an experiment by Creem and Proffitt (2001). They showed that if subjects simultaneously grab objects while trying to recall words (semantic task), they pick up the objects as if blind to the way they are used. The semantic task interferes with grasping objects appropriately by their handles. The task puts heavy demands on the processes needed to retrieve the functional semantics of the object and thus the appropriate grasp to use. Without semantic processing from the vision-for-identification stream, the visuomotor system can still direct the effective grasp of an object, but not in a manner that is appropriate for its use.

Importantly, similar effects are found in driving, for example a meta-analysis of the impact of talking on a mobile phone concluded that reduction in driving performance is primarily on reaction time, not lane keeping performance (Horrey & Wickens, 2004). Thus, talking on the phone may similarly be putting heavy demands on-, or competing with the cognitive processes that are needed for identifying what to do in a particular situation and for setting the action goal of how to react to a stimulus.

Cortical damage in the vision-for-identification stream causes problems with describing the size, shape, or location of objects and ability to imagine shapes and faces, yet amazingly such patients have no problem with locomotion (Kalat, 2004). Milner and Goodale's (1995) famous vision-for-identification-damaged patient DF is able to hike over difficult terrain and avoid obstacles as skillfully as intact individuals (Goodale & Milner, 2004). Locomotor control, the ability to visually control the path taken through the environment and the ability to control collisions, are classic vision-for-action tasks.

Although there may be a "basic protection" by the vision-for-action stream that enables us to act regardless of the semantics of objects, as demonstrated by Creem and Proffitt (2001) or Goodale and Murphy (1997), actions may be improved or augmented by semantic information provided by vision-for-identification. For example, handles are grasped according to their use or reaction time is improved by situation identification and efficient action goal-setting.

The conclusion regarding functional distinction also gains support from the fact that the receptive fields of cells in the vision-for-action stream have a very large representation of the peripheral visual fields. In contrast, cells in the vision-for-identification stream are centered on or around the fovea, the retinal region of highest acuity, and include very little of the far peripheral visual fields (Baizer et al., 1991; Goodale & Humphrey, 1998). Consequentially, the visual control of some motor behavior (e.g. grasp aperture) is equally sensitive to differences in visual stimuli presented in the far peripheral visual field (70 deg) as close to fixation (5 deg), whereas the variability of judgments of the same stimuli increases substantially with eccentricity (Goodale & Murphy, 1997).

However, the fovea is not reserved for the vision-for-identification stream, there is competition for use of the fovea by the vision-for-action stream. Gaze is typically directed to specific locations or "anchor points" that are critical for the planning and control of action (e.g. Flanagan & Johansson, 2003). It follows that if the fovea is not resting on an object or any object conveying "information for identification", like when a driver looks at the road centre or the hiker looks at her path a few steps ahead, then vision is probably primarily carrying information for vision-for-action.

With this as a background we are now better equipped to start thinking about how semi-independent processes interact in driving. The next step is to examine how they are affected by eye movements.

## Vision, attention, and eye movements

We simultaneously attend to and look at objects in a visual scene by means of saccadic eye movements that rapidly bring the fovea onto stimuli of interest. This is called overt attention. It is possible, however, to attend to peripheral objects away from the line of gaze by moving our 'inner eye'. This is called covert attention. Strong evidence indicates that shifting attention to a peripheral location corresponds to a) the oculomotor system preparing to make an eye movement toward it, and b) the preparation of visually guided actions toward that location (Corbetta, Akbudak, Conturo, Snyder, Ollinger, Drury, et al. 2000; Craighero & Rizzolatti, 2005; Rizzolatti et al., 1987).

In direct support of Rizzolatti's premotor theory of attention, Corbetta et al. (2000) discovered, using brain imaging, that attentional and oculomotor processes are tightly integrated at the neural level in overlapping networks. The functional anatomical overlap between attention and eye movements indicates that the same areas are active when people *covertly* attend to peripheral visual stimuli or actually perform directed saccadic eye movements to the same stimuli. So attention and eye movement processes are not only functionally related, they also share anatomical areas in the human brain. Paying attention to an object is clearly linked to eye movements and being able to act on that object.

The fixation act is therefore the primary method of paying attention, but it is supported by covert attention to give a peripheral preview advantage for the next fixation location. Covert attention makes sense only when considered as part of an integrated visual system (Findlay & Gilchrist, 2003). The control of attention may thus have piggy-backed, in evolutionary terms, on the control of eye movements (Goodale & Milner, 2004).

There is however, one important difference between overt and covert attention – overt attention (by fixation) gets the immediate benefit of high-resolution foveal vision and an increasing proportion of cortical processing as the signal proceeds to higher levels of processing (Findlay & Gilchrist, 2003). Similarly, items that are not fixated receive greatly reduced processing. Covert attention does improve the detection and discrimination of stimuli presented at in visual periphery (e.g. Posner, 1980), but the magnitude of the effects of covert attention are relatively small in comparison to overt attention.

Further support for an attention-eye-movement link comes from the striking similarity of eye movements with the strongly supported view that spatial attention can be conceived as a searchlight, spotlight, or zoom-lens (Allport, 1993; Deyoe & Brefczynski, 2005; Mack & Rock, 1998; Neumann &

Sanders, 1996). Spatial orienting of visual attention is accompanied by enhanced neuronal responsiveness and corresponding RT benefits to visual stimuli occurring at the attended location. But spatial attentional engagement also acts as a 'hold' mechanism producing 'fixational capture' that causes a reduction in the responsiveness of saccadic mechanisms to peripheral stimuli and slowed overt responses to stimuli in nonattended locations (Allport, 1993, Rizzolatti et al., 1987).

### Eye movements in driving

Like an advance patrol, eye movements gather preparatory information required for guiding actions. In natural dynamic activities the common finding is that gaze is controlled proactively, not reactively. Proactive gaze lead has been shown to be critical for a range of activities such as object-manipulation (Flanagan & Johansson, 2003; von Hofsten, 2004), table tennis (Land & Furneaux, 1997), reading text and music (Land & Furneaux, 1997), tea-making (Land & Hayhoe, 2001), walking (Patla & Vickers, 1997; Hollands et al., 1995), and driving (Land, 1992; Land & Lee, 1994; Land & Horwood, 1995). In these activities, gaze is typically directed to specific locations or "anchor points" that are critical for the planning and control of an action. Gaze anchor points are actual or potential contact points about 1 second ahead, to which actions are oriented (Johansson et al., 2001; Flanagan & Johansson 2003; Land & Furneaux, 1997; von Hofsten, 2004). According to Flanagan and Johansson (2003) gaze anchor points represent "spatiotemporal checkpoints for the development, maintenance, and adaptation of sensorimotor correlations required by predictive control mechanisms". Thus, the brain prefers proactive, fovea-centered, goal-position information and a "do it where you look" strategy while performing actions (Land & Hayhoe, 2001). The fixation act is not only the most effective mechanism for attention deployment, but also the most effective mechanism for trajectory aiming.

Based on driving performance data, Land and Horwood (1995) determined two distinct regions of information uptake in the visual control of steering: a) a distant region at about 4 degrees down from true horizon (0.93 s or 15.7 m ahead at 61 km/h), and b) near region at about 7 degrees (0.53 s or 9m ahead at 61 km/h). The distant region attracts gaze fixations and is used to match road curvature, whereas the near region is viewed peripherally and is used to keep a proper distance from the lane edges. These results strongly support Donges (1978) two-level control model of steering which uses a predictive (feed-forward) control loop in combination with a near road (feedback) control mechanism. The model implies that, when road curvature has been anticipated by the distant region mechanism, the near region mechanism only needs to fine tune the system. Further supporting Land and Horwood's two-region hypothesis, Salvucci and Gray (2004) recently dem-

onstrated that their two-point control model successfully accounts for curve negotiation with occluded visual regions, corrective steering after a lateral drift, lane changing, and individual differences.

With regard to identifying which distant gaze anchor points are used, Wann and colleagues (Wann & Swapp, 2000; Wann & Wilkie, 2004; Wilkie & Wann, 2002, 2003a, 2003b, 2005) demonstrate that steering becomes highly efficient and simplified when observers fixate the target that steering is directed to. Wann and Land (2000) and Wann and colleagues argue that the most effective anchor point for the distant region is a point on the future path, and that gaze fixation is the main mechanism for trajectory aiming. But other distant region gaze anchor points can be used, including a lead car, which also has the added benefit of providing headway control information (Salvucci, Boer, & Liu, 2001; Crundall, Chapman, Phelps, & Underwood, 2003), the tangent point (Land & Lee, 1994), the vanishing point where the road lines meet in the distance (Salvucci & Gray, 2004), and the focus-of-expansion during relatively straight driving (Gibson, 1958).

Likewise, several sources of near region visual information have been proposed, such as optic- or retinal flow (Lee & Lishman, 1977; Wann & Wilkie, 2004), an estimation of a near road point directly ahead of the vehicle (Salvucci & Gray, 2004), and alignment of bodywork on a car with exterior cues (e.g. alignment of lane markings with the edge of the windshield or a hood ornament) (Wann & Wilkie, 2004).

It is proposed here that the two region hypothesis also gains support from inherent optic flow characteristics. Rotational discrepancy between road optic flow and road geometry is easier to detect in the distant region because near and far points in optic flow contribute differently to the separation of rotational and translational components of optic flow (Grigo & Lappe, 1988). As the rotational component of all flow vectors is identical, the difference between any two flow vectors depends only on the translational component (Lappe et al., 1999). Thus, gazing at the distant region enables the comparison between the rotational component in optic flow (current path) and the road geometry (desired path) at a preview time sufficient for smooth action control. Translation is more prominent in the near region because of motion parallax, whereby near features pass by faster than far features. The flow in the distant region is more stable, and the rotational component is not confounded by translation as the near region. The most salient translation information is near, moving, and better suited for uptake by peripheral vision mechanisms (Leigh & Zee, 1999).

For accurate steering, information from both regions is needed. When distant region information is poor, like driving at night, in fog, or when gazing away from the road, steering becomes increasingly jerky and reliant on other estimates (Land & Horwood, 1995; Wilkie & Wann, 2002). At slower speeds the near road mechanism may be adequate on its own, but at higher speeds it is not (Land & Horwood, 1995). So, although accurate steering

performance is best specified by fovea-centered distant region information, we can still manage somewhat using other estimates. Visual input gracefully degrades in quality and is not all-or-none dependent on the eyes fixating the distant region. Thus, off-road glances give graded amounts of driving information depending on degree of eccentricity. For example, it has been shown that more information for lane-keeping and forward object detection is available from peripheral vision during in-vehicle glances to high and centrally placed displays than glances to low placed displays (Lamble, Laakso, & Summala, 1999; Summala, Nieminen, & Punto, 1996).

### **Time sharing and the two visual streams**

The continuous uptake of information for path- and headway-control in driving has to be satisfied in the presence of other tasks requiring vision, such as checking moving and stationary objects in the visual periphery, reading road signs, and monitoring in-vehicle displays. When these “secondary” tasks require vision, a time sharing behavior is exhibited with the eyes being continuously shifted back and forth between the road center and the off-path object. Importantly, drivers time share not only between the road centre and in-vehicle tasks, but also time share between the road centre and other driving-related objects such as signs, bicyclists, mirrors, scenery and so on (Land, 1998).

Glance-based measures, such as total glance duration, glance frequency, single glance duration, and total task time are the central measures of interest in assessing the visual- or attentional demand of in-vehicle information systems ISO (15007-1). A glance describes the transition to a given area, such as a display, and one or more consecutive fixations on the display until the eyes are moved to a new location ISO (15007-1). The temporal characteristics of glances between road centre and a peripheral object is remarkably constant with glance durations typically exhibiting means between 0.6 to 1.6 seconds, and showing a (positively) skewed distribution towards short glances (Wierwille, 1993a; Green, 1999). For conventional instrument panel functions such as the speedometer, radio, clock, etc, the longest mean single glance durations range from 1.2 to 1.85 seconds depending on the study. There is a large consensus that drivers generally are very unwilling to look away from the road for more than 2 seconds (e.g. Green, 1999; Rockwell, 1972).

Recall that a large amount of converging evidence shows that sensorimotor representation can only be expressed within a short delay of about 2 seconds following stimulus presentation (Rosetti & Pisella, 2002). This dramatic effect of delay on sensorimotor processes is of tremendous significance for explaining the time sharing regularity of off-road glances. A glance away from the road disrupts the more spatially accurate vision-for-action information and leaves us increasingly dependent on information that was originally delivered by vision-for-identification. However, some online

near region information is accessible by peripheral vision during off-road glances of low visual angle. It is likely that higher weightings are given to this peripherally-derived type of information when looking away, but it has been clearly shown to be a sub-optimal strategy requiring more attention (Wann, Swapp, and Rushton, 2000). See Wilkie and Wann's model (2005) for an example of how cues could be weighted in different situations.

### **The gaze concentration effect - adaptive visual guidance behavior in path control**

Evidence of the importance of the road centre region for action guidance comes from the way eye-movement behaviour adapts to action task demands. Concentration of gaze to the road centre area is the driver's behavioral response to increased demands caused by driving task complexity, secondary task demands, and driver states.

Consider the way a rally driver has his/her eyes 'glued' to the road ahead, only rarely looking away. The action task is so extreme that a co-driver is needed to assist with planning and recognition tasks that would require removing the eyes from the road. This observation is supported by research showing that drivers devote more attention to the control of the vehicle by looking more at the road centre as the driving task becomes more difficult (Muir, 1990; Senders, Kristofferson, Levison, Dietrich, & Ward, 1967). In curve negotiation, gaze spatially concentrates more to the road centre region than it does in straight driving (Land & Lee, 1994; Wann & Swapp, 2000). On-road occlusion studies (Senders, et al., 1967), where vision has been blocked momentarily, demonstrate that drivers dramatically decrease eyes-off-road-times as speed is increased. (Senders, et al., 1967). Also, Land and Horwood (1995) indicate that drivers also look further up the road when driving faster.

During visual tasks, road centre viewing time is dramatically disrupted and reduced because of in-vehicle glances. When visual task difficulty increases, drivers look less at the road ahead because they look more often, for longer periods, and for more varied durations at the in-vehicle display (Green, 1999). However, when drivers are faced with increasing *driving* task demand during the performance of in-vehicle visual tasks, they adapt their glance behavior by increasing viewing time to the road or by slowing down (Green, 1999; Lansdown, 2001; Rockwell, 1972; Senders et al., 1967; Wierwille, 1993b), thus compensating for both anticipated and unanticipated increases in driving task demands (Wierwille, Antin, Dingus, & Hulse, 1988).

In response to increases in cognitive or auditory task difficulty, drivers increase their road centre viewing time and spatially concentrate their gaze in the road centre region at the expense of peripheral glances. This gaze concentration effect has reliably been found for cognitive and auditory tasks (Hammel et al., 2002; Harbluk et al., 2002; Nunes & Recarte, 2002; Recarte

& Nunes, 2000, 2003). Significant reductions in horizontal and vertical variability (SD) of gaze direction, longer on-road fixations (more staring), and reduced glance frequency to mirrors and speedometer are typically found.

Impairment of driver state has also been associated with gaze concentration, as with alcohol (Belt, 1969), anxiety (Janelle et al., 1999) and fatigue (Kaluger & Smith, 1970). For example, Belt (1969) found a significant concentration of eye movement patterns at a blood alcohol level of 0.08 mg/cm<sup>2</sup>. Drivers elicited almost a complete lack of search outside a 3° x 3° space near the focus of expansion, e.g. passing vehicles were always fixated in the control condition and not at all at the 0.08 level.

The amount of time spent looking at road centre has a strong relationship with crashes and driving performance measures such as path control and reaction time measures. The recent real-world crash statistics from the 100-car study (Neale, et al., 2005) point to the overwhelming importance of attention to the forward roadway for crashes and near miss events. In line with this, Wierwille and Tijerina (1988) show high correlations of  $r=.90$  between eyes off road exposure (mean glance duration \* number of glances \* frequency of use) with real-life crash frequency.

It is possible that real-life crash data points to a stronger relationship between eye movement measures and crashes than lane keeping measures and brake reaction time measures. Correlations between eye movement and lane-keeping measures are typically in the  $r=.60$  to  $.80$  range (Ito & Miki, 1997; Green, 1999; Wierwille, et al., 1988; Zhang & Smith, 2004a). Zhang and Smith (2004b) show correlations between measures of off road glances and brake reaction time of  $r=.65$ . It is proposed here that the reason for the discrepancy between real-world and experimental results has to do with the degree expectancy of the events used in these experiments. In the 100-car study, drivers were found to have difficulty responding appropriately when other vehicles performed unexpected maneuvers and when expectancies about the flow of traffic were violated in conjunction inattention (Neale, et al., 2005). Also, Green (2000) found in a comprehensive review that the most important variable affecting brake reaction time is driver expectation. Surprise intrusions give much longer reaction times (1.5 s) than unexpected (1.25 s), and expected signals (0.75 s). Thus, eyes-off-road is likely to be higher correlated with reactions to the unexpected braking and lane keeping events that are more involved in crashes.

During cognitive tasks, event reaction time deteriorates, but there is little interference with path control. Recarte and Nunes (2000, 2003) have shown that the gaze concentration effect in cognitive tasks is associated with loss of event detection capability across the entire visual field. This is also supported by Harbluk et al's (2002) finding of more incidents of hard braking during cognitive tasks. Similarly, in a meta-analysis of the impact of talking on a mobile phone, Horrey and Wickens (2004) conclude that reduction in driving performance is primarily on reaction time, not tracking performance.

Taken together, these findings indicate that visual guidance of path-control is more immune to interference from cognitive tasks than identification and planning tasks. The main conclusion that can be made from the way gaze concentrates in different demanding situations is that the road centre area is of central importance for driving performance.

### **Gaze concentration and ‘tunnel vision’**

Gaze concentration and reduced sensitivity to stimuli in the visual field (variously referred to as tunnel vision or cognitive tunnelling in the useful field of view, functional field of view, or visual lobe) are not identical, but they do seem to occur simultaneously (e.g. Recarte and Nunes, 2003) and are frequently confused. Gaze concentration refers to the reduced spread of fixations over time, whereas tunnel vision refers to reduced detection performance across the visual field during a fixation (independent of eye movement). The common basic finding with regard to reduced visual field sensitivity effects is that higher processing demands, such as those induced by cognitive tasks, driving environment demands, anxiety, and vigilance, produce performance deterioration within the visual field. One area of disagreement regarding visual field sensitivity reduction has been whether the sensitivity deteriorates more strongly in peripheral regions or not. It appears that the general interference hypothesis has much clearer support (e.g. Crundall, et al., 2002; Recarte and Nunes, 2003), see Figure 1. This general interference hypothesis is also supported by an important finding by Mack and Rock (1998), wherein almost all subjects failed to see an unexpected object appearing right at the fovea while they performed a discrimination task covertly in visual periphery.

## **Attention**

Although, James (1890) famously exclaimed “everyone knows what attention is”, only six years later Groos (1896) stated “To the question, ‘What is Attention?’ there is not only no generally recognized answer, but the different attempts at a solution even diverge in the most disturbing manner”. Surprisingly, Groos sums up the recent situation quite well (see Tsotsos, Itti, & Rees, 2005).

There is an enormous amount of literature on attention, yet it is a notoriously slippery concept, difficult to precisely pin down. One persistent problem is with the definition of attention. Attention is a vague, catchall phrase, permitting a great variety of meanings, such as arousal, alertness, activation, awareness, consciousness, effort, capacity, resource, etc. In the accident research above, inattention is referred to as drowsiness, distraction, looking away from the road ahead, and being engaged in thought. Importantly, the variety of meanings also illustrates that it is extremely difficult to study at-

tention without confounding it with other issues of perceptual, motor, or central processing. Attention cannot be studied in isolation; it always accompanies some other process (Neumann & Sanders, 1996).

In spite of all the conflicting theories and negative viewpoints throughout the years, attention is now one of the central topics of modern neuroscience research thanks to the availability of new experimental techniques, such as functional brain imaging (see Itti, Rees, and Tsotsos, 2005). As we shall see, these new techniques have recently led to a far clearer picture of attention.

### Attention is selection

Attention is often operationalized as a selection mechanism for relevant information (Itti, Rees, and Tsotsos, 2005; Neumann & Sanders, 1996). For example, a commonly agreeable basic statement describing attention is “Key to the survival of many biological organisms is their ability to selectively focus neural processing resources onto the most relevant subsets of all available sensory inputs.” (Tsotsos, Itti, & Rees, 2005, p. xxi). One exemplary definition, applied to vision, comes from Rensink (2002b): “attention is simply the control of information in the visual system, carried out by a set of selective processes that can be co-ordinated via their operations on a target structure”. In this thesis, focus is primarily on visual attention, secondarily on cognitive, and auditory attention, and not at all on driver impairment related causes such as drowsiness.

Current theories tend to emphasize an active selection process rather than a passive ‘protective’ process filtering out unwanted information. This emphasis sees selection as a result of the organism striving to be efficient rather than trying to protect itself from a bombardment of information. For example, the use of a reduced representation of the environment or a subset of information, allows computations to be performed faster and more efficiently when under time pressure as compared to an approach in which the entire input is represented (Billock, Koch, & Psaltis, 2005). Limits of attention are thus a consequence of the way in which the brain solves selection problems in the control of action (Neumann, 1987).

Most current models of attention argue that attention biases competition for neural representation either in favor of attended stimuli, against unattended stimuli, or some combination of both (e.g. Bundesen, Habekost, & Kyllingsbæk, 2005; Logan, 2002; Somers & McMains, 2005; McAdams and Maunsell, 1999). Attention can thus be seen as a selection process where some inputs are processed faster, better, or deeper than others, so that they have a better chance of producing or influencing a behavioural response, for example by speeding up reaction time (adapted from Lamme, 2005).

## Stimulus-driven attention and goal-directed attention

At one extreme, attention can be dominated by external events, being drawn to salient objects or distracting events. On the other extreme attention is controlled by cognitive factors such as knowledge, expectation, and current goals. Corbetta and Shulman (2002), and Koch (2004) show quite conclusively that visual attention is controlled by two functionally and neurologically distinct neural systems a) a bottom-up, transient, stimulus-driven, saliency-based attention system and b) a top-down, sustained, goal-directed attention system. This rejects the most deep-seated assumption in traditional views, according to Allport (1993), that there exists a unique and unitary central system (or attentional system, or central executive) of limited capacity, that can only be bypassed by (unlimited) automatic processes.

Stimulus-driven, bottom-up attention is driven by intrinsic stimulus qualities, acts rapidly, acts automatically, mediates pop-out, and acts across the entire visual field (Corbetta & Shulman, 2002). Stimulus-driven attention is specialized for the detection of behaviorally relevant stimuli and acts as a 'circuit breaker' for goal-directed attention, directing attention to salient events (Corbetta & Shulman, 2002), see Table 1. Selection is believed to be controlled via one or more explicit saliency maps, wherein neurons encode saliency or conspicuity of objects in the visual environment, not stimulus attributes (Corbetta & Shulman, 2002; Itti, 2005; Itti & Koch, 2000; Koch & Ullman, 1985). Competition among neurons in a saliency map gives rise to a single winning location that corresponds to the next attended target.

Goal-directed, top-down attention is involved in the cognitive selection of sensory information and responses, takes longer to deploy and can be directed to either a proscribed region in space, to individual objects, or to specific attributes throughout the visual field (Corbetta & Shulman, 2002; Koch, 2004). Determining which stimulus in a visual display is the most important to an observer requires the integration of both top-down and bottom-up processes. Thus, stimulus-driven aspects interact with ongoing cognitive goals (Corbetta & Shulman, 2002). Attentional selection is performed based on top-down weighting of the bottom-up feature maps that are relevant to a target of interest (Itti, 2005). This dynamic interaction is central to current theories of attention such as biased competition accounts (Desimone & Duncan, 1995; Itti, 2005; Reynolds & Desimone, 1999) and similar elaborations (Bundesen, Habekost, & Kyllingsbæk, 2005; Logan, 2002). The dynamic interaction explains how certain stimuli are enhanced, while other unattended stimuli are suppressed (see Koch, 2004; McAdams & Maunsell, 1999). In short, attention is the biased sum of bottom-up and top-down processes when it serves visual identification and selection of objects.

### **Change blindness and Inattention blindness exemplify the interaction between stimulus-driven and goal-directed attention**

Almost everyone has had the experience of looking without seeing something. For example, most drivers have experienced brief moments of not seeing or “sighted blindness”, which produce astonishment and alarm when awareness returns.

The key result of the change- and inattention blindness paradigms is that absence of attention causes apparent blindness (Mack & Rock, 1998; Rensink, 2002a). Detection of a change is dependent on either a) visual transients alerting to a change (i.e. stimulus-driven attention) or b) a representation of the feature that changes (i.e. goal-directed attention). The visual system’s ability to respond is severely impaired when visual transients, a kind of low-level feature detection mechanism (Stelmach et al., 1984), are masked by common visual disruptions such as saccades, blinks, mud splashes, and occlusions (Rensink, 2002a). If a change occurs simultaneously with a visual disruption that wipes out the attention-grabbing visual transients, then change detection relies on there being a representation of the changing feature for detection to occur. That is, if we are ‘keeping an eye on something’ or can expect what will be changed and therefore have our attention directed to it, we will see the change. However, if the change is unexpected, and especially if visual transients are masked or disregarded, it is very unlikely to be detected.

The concept of inattention blindness emerged from a recurrent finding that about 25% of subjects fail to detect stimuli when they are unexpected and attention is directed elsewhere (Mack & Rock, 1998). Detection rates vary systematically as a function of the degree of similarity between an unexpected object and the attended items (Ambinder & Simons, 2005; Most, Scholl, Clifford, & Simons, 2005; Simons & Chabris, 1999). Thus, lack of expectation alone, without visual transient disruption, causes apparent blindness.

### **Action-driven attention**

Most attention research focuses on how (stimulus-driven- and goal-directed) attention serves the conscious visual identification and selection of objects in experiments involving passive observation, thus focusing on vision-for-identification processing (see Bundesen, Habekost, & Kyllingsbæk, 2005). It is very important not to forget that attention also has to serve the requirements of actions. A different, action-driven selection of visual input is needed for the control of action. Action-oriented approaches to attention are reviewed in Allport (1987, 1993), Neumann (1990), and Neumann and Sanders (1996). Similar concepts to action-driven attention have been variously named *selection for action* (Allport, 1989; Neumann, 1987), “selection

of the stimuli that control an action” or “unconscious encoding of sensory stimulation” (Mack & Rock, 1998), ‘*zombie*’ vision (Koch, 2004), and vision-for-action (Milner & Goodale, 1995).

Unfortunately, a long-standing tradition in psychology implicitly equates attention with conscious awareness. In fact, a clear distinction can be made between attention and conscious awareness, especially from the cognitive neuroscience perspective (Lamme, 2005). For example, the stimulus-driven selection process (above) is not under conscious control, although stimulus-driven- and goal-directed attention both lead to explicit awareness (Corbetta & Shulman, 2002; Koch, 2004). The dependency of attention on visual transients, as in the change blindness research, also points out the importance of unconscious stimulus-driven attention for ‘triggering’ explicit awareness. See Table 1 for a distinction of the various forms of attentional selection.

Table 1. Three forms of attentional selection (adapted from Koch, 2004).

Property	Action-driven attention	Stimulus-driven attention	Goal-directed attention
Function	Action control	Both action control, and identification and cognition	Identification and cognition
Under conscious control	No	No	Yes
Facilitates explicit awareness	No	Yes	Yes
Spatial specificity	Throughout entire visual field	Throughout entire visual field	Spatially-limited
Feature specificity	Dependent on what an action requires	Acts at all times and in all feature dimensions (Saliency)	Can select specific attribute
Duration	Transient	Transient	Sustained (with effort)
Task dependency	Yes	No	Yes

The two-stream vision section above established that what we think we ‘see’ is not always what guides our actions. The vision-for-action visuomotor modules are able to attend to or select the visually-derived information they need to perform certain actions such as locomotion without involving awareness. Therefore, in line with the premotor theory of attention (Rizzolatti et al., 1987) and Atkinson (2000), action-driven attention is considered here to be identical to the selection processes that are involved in the vision-for-action stream. Action-driven attention involves activity in a number of sensorimotor action modules such as oculomotor and those for reaching and grasping and locomotion (Rizzolatti et al., 1987). Attentional modulation in vision-for-action, noncategorical, spatial-vision systems entails the temporary decoupling of unattended visual locations from potential motor command, which costs reaction time delays that reflect the cost of disengagement from attended location. (Allport, 1993).

Action-driven attention is perhaps best illustrated by the way vision-for-identification damaged patients have no problem with locomotion or grasping although they cannot *consciously* attend to the size, shape, or location of objects (Kalat, 2004; Milner & Goodale, 1995). Other prominent examples of the dissociation between explicit awareness and information selection in visuomotor systems include the oculomotor system being able to compensate for a consciously invisible shift of target position during a saccade (Goodale, Pélisson, & Prablanc, 1986), that subjects are unaware of their body posture adjusting to a swinging room (Lee & Lishman, 1977), and that subjects are able to correctly estimate the steepness of hills by tilting their hand although verbal and visual ratings consistently overestimate slope (Proffitt, Bhalla, Gossweiler, & Midgett, 1995).

This distinction is necessary to explain situations where explicit conscious awareness fails under conditions of inattention, as in inattention blindness, yet action-driven attention and automatic visuomotor processes (zombie behaviors) do not. When we continue to drive during distracting situations without explicit conscious awareness of driving, our sensorimotor processing is obviously getting its information from somewhere.

Although peripheral vision plays a substantial role, the main mechanism of information selection for action is foveal vision. Fixations are very tightly determined by the needs of the action task but yet we remain unaware of many aspects of eye movements. This unawareness of how we use our eye movements to guide action has led to a string of ‘discoveries’ of how we fixate our environment during a variety of actions, such as driving (Land, 1992; Land & Lee, 1994; Land & Horwood, 1995), walking (Patla & Vickers, 1997; Hollands et al, 1995), object-manipulation (Flanagan & Johansson, 2003; von Hofsten, 2004), table tennis (Land & Furneaux, 1997), reading text and music (Land & Furneaux, 1997), and tea-making (Land & Hayhoe, 2001), etc. The mere fact that we discover the way we use our eyes is in itself evidence that these processes are not entirely available to introspection.

The main conclusion regarding attention is that we have to serve the needs of all three simultaneously active attentional processes: 1) a process serving the on-line information needs of actions, 2) a bottom-up, saliency driven process that is on the look-out for changes and salient features in the environment, drawing attention to them when they reach a threshold, and 3) a process that is closest to cognition and consciousness, directing and spreading out attention at will.

## Prediction – Expectation and Extrapolation

To understand crashes it is essential to understand expectancy. Recall that one of two key causal factors in crashes is the occurrence an unexpected event. Expectancy is a form of prediction, meaning to await, to anticipate, or

consider probable *the occurrence* of something. Thus, it generally refers to more discrete events. Whereas, to extrapolate is to predict by *projecting past data* or experience, and thus extrapolation deals with more continuous data.

Accordingly, prediction involves a combination of two basic types of influences on behavior a) non-linear *decision-based* prediction influences which characterize discrete, inherently nonlinear aspects of prediction, and b) linear *trajectory-extrapolation-based* prediction influences which represent continuous low-level processes (see Kowler, 1990; Maciejowski, 2002; Pavel, 1990; von Hofsten, 1993, 1995; and Wolpert, et al., 1998). In driving, continuous trajectory extrapolations (e.g. vehicle paths) occur simultaneously with discrete expectations and decisions (e.g. traffic rules).

A key point is that we act upon *predicted* future states of our body and the future of objects and events evolving in the world to act smoothly and efficiently, *not* immediate visual input (Gredebäck, 2004; Kowler, 1990; Pavel, 1990; von Hofsten, 2004). Lags in biological feedback loops imply that *all* visually guided actions require some degree of prediction and have to be directed towards future states (von Hofsten, 1993, 1995, 2004). Prediction is both embedded in our online sensorimotor systems in the form of extrapolation influences (Blakemore et al. 1999; Weiskrantz et al, 1971; Wolpert, et al., 1998) and in higher cognitive functions as decision-based influences (e.g. Ericsson & Smith, 1991; Mack & Rock, 1998; Pavel, 1990). Importantly, these predictive requirements cause eye movements to be proactive during action.

There is abundant evidence that humans are very well equipped to perceive extrapolations such as time-to-collision, time-to-object, and time-to-line-crossing (Field & Wann, 2005; Nilsson, 2001). These time-related variables, and similar concepts such as the safety margin (Näätänen & Summala, 1976) and the field-of-safe-travel (Gibson & Crooks, 1938), specify future states and thus are excellent examples of the use of trajectory-based prediction to represent an area of predicted safe driving around the vehicle.

Although the control of time-to-collision and path control are classic vision-for-action tasks (Field & Wann, 2005), expectation has a large impact on reaction time (Corbetta & Shulman, 2002; Evans, 2004; Green, 2000; Neumann & Sanders, 1996). Many reactions also rely on understanding signs and symbols involving vision-for-identification and higher cognitive processing, for example braking associated with an exit sign. Behavior may depend on the expectations specified by the prior probabilities of events, for example a vehicle running a red light. Thus, predictive effects due to decision-based prediction complement the effects of extrapolation.

# Empirical studies

## General aims

Against this theoretical backdrop the results of empirical studies will be examined. As mentioned, the second aim of this thesis is to apply this theoretical knowledge to the results from the empirical studies in order to explain how visual performance is affected by various activities while driving.

The studies presented here were performed within the constraints and context of a research project called Driving Support from VISual Behavior RECognition (VISREC), cofunded by the Program Board for Swedish Automotive Research and Volvo Technology. The basic objective of the VISREC project was to prevent accidents due to lack of attention by envisioning and implementing real-time, in-vehicle attention support functions. The VISREC project work was carried out in three areas – a) sensor development, b) inattention detection, and c) attention support functionality development. Study III was also carried out in collaboration with Transport Canada (and their subcontractor the University of Calgary), and the Swedish National Road and Transport Research Institute (VTI) within the EU project Human Machine Interface And the Safety of Traffic in Europe (HASTE).

Before this thesis was started in 1999 it was notoriously difficult and time consuming to measure and analyze eye movements in an in-vehicle setting outdoors in an automatic manner. A precondition for studying inattention through eye movements in driving on more than just a few subjects was therefore to develop both a sensor that would be robust enough to study in-vehicle eye movements and an analysis method to eliminate manual analyses. The aims of study I were a) to present the result of the development of a suitable eye-tracking sensor, b) to validate an eye-tracking analysis method, and c) to examine the visual demands of five common in-vehicle tasks in comparison to reading an SMS text message on a mobile phone.

Study II put the eye-tracker and the analysis method developed in study I to use. It was aimed at examining how driving information acquisition is affected by everyday in-vehicle tasks during naturalistic driving. The main hypothesis was that the road center region carries the most important visuo-motor information for driving and therefore is prioritized while in-vehicle visual and cognitive tasks are performed. New eye movement measures were

developed to enable a comparison between the overall effects of visual tasks, cognitive tasks, and baseline driving. Additionally the specific effects of the following factors on eye movement measures were studied: a) display position, b) hands-free phone installation, c) dialing in comparison to other commonly acceptable tasks, d) new information system tasks (such as navigation system tasks or reading emails), e) conversation medium, and f) cognitive task type.

Study III was aimed at answering four research questions. How are eye movements influenced by (1) different in-vehicle task types (visual and auditory), (2) increasing in-vehicle task difficulty while controlling for task length, and (3) driving task complexity? Lastly, (4) which measures are most suitable and sensitive to these changes in eye movements?

## Automated and validated visual behavior measurement (Study I)

Study I presents a) the result of the development of a suitable eye-tracking sensor, b) the results of a validation of an offline automated analysis algorithm that eliminates the video transcription process commonly used to measure visual demand (ISO 15007-1, ISO 15007-2), and c) the visual demands of five common in-vehicle tasks (looking in the mirror, reading text to the left and right in the instrument cluster, adjusting the fan, and adjusting controls on the radio in comparison to reading an SMS text message on a mobile phone. The most basic aim of this study was to validate this automated analysis algorithm as compared to a standardised video transcription based method.

The new eye-tracking sensor, called faceLAB, was the result of research collaboration between the Australian National University and Volvo Technology under the author's responsibility for the purposes of this thesis' and the VISREC project. Volvo and the Australian National University later decided to invest their rights in the technology into a spin-off company called Seeing Machines and to turn faceLAB into a commercial product. Studies I, II, and III all used versions of this eye-tracker.

A standardized in-vehicle visual demand measurement method (ISO 15007-1, ISO15007-2) describes procedures and definitions of variables for visual demand measurement. Four glance-based measures are the central measures of interest in assessing the visual- or attentional demand of in-vehicle information systems – total glance duration, glance frequency, single glance duration, and total task duration (ISO 15007-1), these measures are illustrated in Figure 3 and described in Table 2. The four measures validated here are defined together with all eye-movement dependent variables used in

studies II and III in Table 2. Video transcription was done in conformance with the ISO method (ISO 15007-1, ISO 15007-2) using the Observer software to provide manually transcribed glance data.

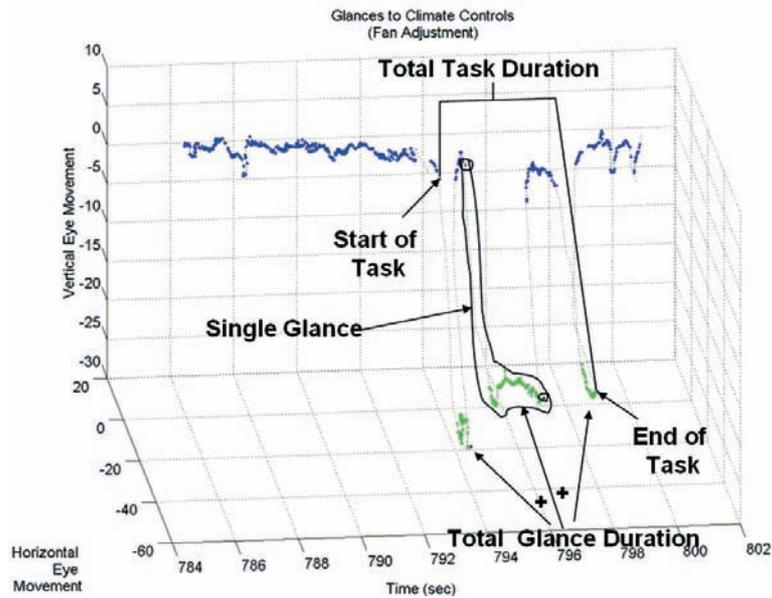


Figure 3. Three glances to the climate controls illustrating the central concepts in the definitions of glance-based measures. Eye movement in degrees.

Table 2. Dependent measures and definitions used in studies I, II, and III.

Measure Type and Name	Abbreviation	Explanation
<i>Glance measures</i>		
Glance frequency	GF	The total number of glances made to a task-related object during a visual task, where each glance is separated by at least one glance to a different target. A glance includes transition time (e.g. one or more saccades) to an area of interest (e.g. a task-related object) and any series of fixations on that area.
Total glance duration	TGD	The total amount of time which glances are associated with visual task related objects.
Total task duration	TTD	The time from the initiation the first saccade to a visual task object to the end of the last glance on a task object.
Mean glance duration	MGD	The mean time of each single glance which start the moment at which the direction of gaze moves towards a task-related object and end at the moment gaze moves away from it.
Standard deviation of glance duration	SDGD	Standard deviation of glance durations.
Number of glance durations exceeding two seconds	G>2S	The number of glances toward the task-related objects during a visual task that have a duration longer than two seconds.
Percent glance durations exceeding 2 seconds	PGD>2	The percentage of glances toward the task-related objects during a visual task that have a duration longer than two seconds.
Glance frequency per minute	GF per min	Rate per minute, calculated by dividing GF by TTD and multiplying by 60 seconds.
Total glance duration per minute	TGD per min	Rate per minute, calculated by dividing TGD by TTD and multiplying by 60 seconds.
<i>Percentage gaze measures</i>		
Percent road centre	PRC	The percentage of gaze data points labelled as fixations during a fixed period of time (e.g. one minute) that fall within a road centre area. The road centre area is defined as a circular area of 16 degrees diameter, centred around the road centre point. The road centre point was determined as the mode, or most frequent gaze angle, of each subject's baseline driving data. The mode was calculated by binning the data in 128 by 128 bins for a 120 by 120 degree portion of the data in the forward view. PRC, when calculated for visual tasks, includes data from the visual task, lasting x seconds, and is filled out with the PRC mean of baseline driving for the rest of fixed period of time, lasting <i>fixed_period-x</i> seconds. An slightly different calculation using a 20x15 degree road center area centred around the subjects mean fixation point in the forward view was also used in study III.
Percent road centre during a visual task	PRCtask	PRCtask is calculated in the same general manner as above, except that it only uses data collected during the duration a task, (i.e. without the PRC baseline driving fill of <i>fixed_period-x</i> seconds). The main difference is thus that PRCtask is not calculated using a fixed period of time, but rather using a variable length of time dependent on each single task duration for each subject.
<i>Gaze variation measures</i>		
Standard deviation of horizontal gaze	SDHG	Standard deviation of the horizontal gaze signal.
Standard deviation of vertical gaze	SDVG	Standard deviation of the vertical gaze signal.
Standard deviation of radial gaze	SDRG/SDG	Standard deviation of the vector sum of horizontal and vertical gaze components (i.e. the square root of the sum of squared vertical and squared horizontal angles), and thus is a one-dimensional angle between the zero intercept and gaze point. In study III two versions are calculated, one using gaze angles and one using distance in cm after gaze is projected onto a forward plane.
Standard deviation of radial external gaze	SDREG	Standard deviation of radial gaze when glances toward a visual task are removed from the data, as such it is only applicable to visual task data

## The initial eye-movement analysis algorithm

An offline analysis algorithm was developed in study I to identify and calculate glances, transitions, fixations, and saccades. A median filter using a 13 sample, 200 ms moving window was developed. The median filter is a non-linear data smoothing technique useful for noise suppression, and is considered better than classical smoothing procedures as it is effective in preserving sharp edges and smoothing spiky noise (Gu, et. al., 2000). The 200 ms window provided the best performance as a trade off between output variance and window length. A fixation identification algorithm was likewise developed. A velocity-threshold based identification algorithm was chosen after comparison with a dispersion-threshold algorithm (see Salvucci & Goldberg, 2000 for an overview of fixation algorithms). The velocity threshold was set at 0.05 rad/frame (172°/s at 60Hz). Data samples above this threshold and two samples preceding and following were removed (to provide an approximation of saccade starts/ends and thus also removing the lower saccadic velocities towards the start/end of the saccade). A glance classification algorithm was also developed to comply with the ISO standard. It identified the starts of transitions and the ends of dwell times (Figure 4). A manual delimitation of clusters was employed to associate glance clusters to a particular area of interest. Statistics on single glance duration, glance frequency, total glance time, and total task time were automatically calculated.

## Results

Pearson product-moment correlations revealed very high correlations between analysis type (video transcription vs. automatic analysis) on all dependent variables (task length  $r=0.991$ , total glance time  $r=0.995$ , glance frequency  $r=0.997$ , and glance duration  $r=0.732$ ). Furthermore, analyses of variance showed no significant differences between analysis types on dependent measures. Figure 6 plots the means and standard deviations for each task, see Figure 4.

Taking the automated analysis data results as a whole, significant differences were found ( $df=11$ ,  $p<0.001$ ) between tasks on task duration ( $F=18.13$ ,  $Mse=20.91$ ), total glance duration ( $F=13.08$ ,  $Mse=10.44$ ), and glance frequency ( $F=28.69$ ,  $Mse=2.94$ ) using an ANOVA. Post hoc comparison with Tukeys HSD showed that it was the mobile phone task that was significantly ( $p<0.001$ ) different from all other tasks. No significant differences were noted between other tasks.

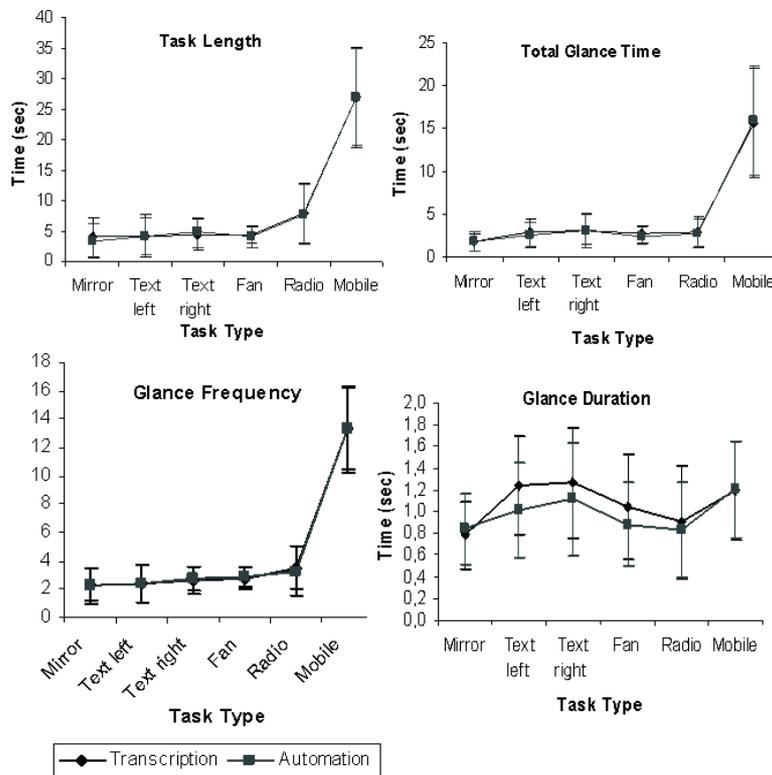


Figure 4. Plots of means and SD for analysis type and dependent variables.

## Discussion

The strong correlations between analysis methods indicate that the automated analysis is indeed valid. The results also showed that reading the SMS message on a mobile phone while driving is judged unsafe by common safety criteria and that it was significantly different than common in-vehicle tasks. The mobile phone task, reading an SMS message, caused long mean task lengths of about 27 seconds compared with roughly 6 seconds for the other tasks. Mean total glance time (time spent looking at the display) was about 15 seconds compared to about 3 seconds for the others. Mean glance frequency was about 13 glances per task compared to about 2.5 glances per task for the other tasks.

The eye movement analysis method presented in study I was later named the Visual Demand Measurement (VDM) tool and analysis method. It was further developed using more sophisticated signal processing techniques and validated once again yielding similar results by Larsson (2003) under the author's supervision for the VISREC project. In conjunction with this work,

the algorithm was somewhat further developed and disclosed in a patent application (Victor & Larsson, 2004), mainly because it also was implemented as an on-line version. The VDM tool, was used in Studies II and III and is described in detail in Larsson (2003) and Victor and Larsson (2004).

The VDM tool takes the raw gaze data produced by faceLAB as input. All gaze data is first transformed to a head normal position, to standardize the visual angle to objects, and all calculations thereafter are based on degrees of driver visual angle. Next, signal-processing algorithms for noise reduction and data quality management are applied to the gaze signal. Special attention was paid to developing methods to identify and either reconstruct or remove poor quality data. Thereafter the signal is segmented into saccades, fixations, and periods of non-tracking using a hybrid algorithm that combines dispersion-, velocity- and eye-physiology-based rules depending on level of data-quality. For the gaze-based measures (see Table 2) in Study II and III, calculations were based on data marked as fixations. Glances are then segmented using an automatic, bottom-up, data-driven clustering of fixations. Thus, glance segmentation does not rely on the gaze to fall within a pre-defined area-of-interest, but rather on fixations being clustered together and only thereafter becoming associated with objects. The glance-based measures used the glance as the unit of calculation. The VDM tool and analysis method is currently being developed further within the EU project Adaptive Integrated Driver-vehicle Interface (AIDE).

As disclosed in Victor and Larsson (2004), this algorithm was developed into an *on-line* version that, when combined with an on-line version of Percent Road Centre algorithm developed in study II, was used in the Distraction Alert function as described below and in Victor (2003).

By eliminating time consuming video transcription, safety testing of in-vehicle information systems tests became easier to carry out, both in fast paced development projects and in general research, thereby removing the main difficulty in using the eye-glance-based method. Further support for this conclusion comes from successful experiences of using of the VDM tool in studies II and III and in various projects, such as HASTE, AIDE, and VISREC.

## Gaze concentration and other eye movement effects of naturalistic tasks (Study II)

Study II aimed to examine how driving information acquisition is affected by everyday in-vehicle tasks during naturalistic driving. The main hypothesis was that the road center region carries the most important visuomotor information for driving and therefore is prioritized while in-vehicle visual and cognitive tasks are performed. Viewing time in the road centre region is

expected to be maximized during task performance, and gaze is expected to be more densely concentrated around a road center point, both at the expense of viewing other areas. It follows that road center measures should be more sensitive than off-road glance measures because they measure more relevant information for visuomotor control in driving.

In addition to the general effects regarding the significance of the road centre region, this experiment is designed to look at certain specific visual and cognitive task effects. The following comparisons are examined. To study the effect of display position on eye movement measures, the same email reading task is included in three different positions (handheld-, radio-, and high navigation system positions). To examine if there is an added visual benefit from phone installation, dialing on a handheld telephone is compared with dialing on a mounted hands free telephone. For comparison purposes, the effect of dialing is also assessed relative to other commonly acceptable tasks, such as radio tuning or resetting fuel consumption. The relative effects of a number of newer tasks requiring a large number of interactions are assessed, such as navigation system tasks or reading an email on a handheld computer. The relative effects of conversation medium on eye movement measures are assessed for question-and-answer conversations taking place in three mediums (through a hand held telephone, through a hands free telephone, and with a passenger). To explore potential differences between a variety of cognitive tasks, listening to an email being read by computerized text-to-speech is compared to a backward counting task, and the three conversation conditions. This work is also aimed at developing eye movement measures to enable a comparison between visual tasks, cognitive tasks, and baseline driving, something which is lacking today.

## Design

16 drivers performed the 18 naturalistic visual and cognitive tasks outlined in Table 3 on a motorway in real traffic. The major differentiating factor between visual and cognitive task types was whether or not the in-vehicle task requires vision to support its execution. See Table 2 for descriptions of the dependent measures used. The PRC, PRCTask, SDRG and SDREG measures were developed during the analysis of the eye movement data in this study. PRC was calculated using a one minute fixed period, see Table 2 for details. PRCTask was calculated for however long it took for each individual task to be completed, the calculation of duration of each task is therefore identical to that calculated for the Total task duration measure, see Table 2 and Figure 3.

The driving task was always present in the experiment and constituted a no-task control condition which could be used to evaluate the effects of 18 different experimenter-requested secondary tasks on various measures of eye movement. They were designed to be a representative sample of everyday

in-vehicle tasks, varying in difficulty, length, and in mixture of visual, manual, auditory, verbal, and cognitive components, see Table 3.

*Table 3.* A description of the secondary tasks used during the experiment.

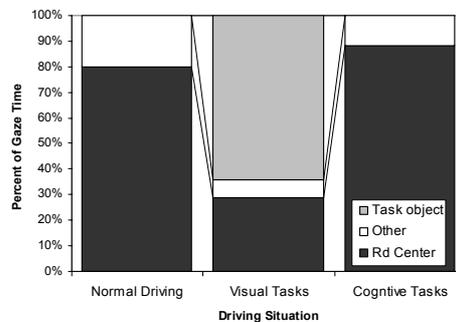
<b>Task Name</b>	<b>Task instruction</b>	<b>Task type</b>
1 Climate	Please raise the temperature and the fan.	Visual
2 Engine Temp	How warm is the engine on the gague in the intrument panel?	Visual
3 CD Track	Please put on song 7 on the CD (starting from radio mode).	Visual
4 Radio Stn	Please find radio station NRJ 105,3 FM (starting from P3 98.4 FM, using the "seek" button).	Visual
5 Reset Fuel Cons	Please reset average fuel consumption on the trip computer.	Visual
6 Coin	Please take out 7.50 SEK from the compartment in the dash (to the left of the radio).	Visual
7 Dial Handheld	Please dial number 031-66 51 81 on the handheld phone and await a set of questions.	Visual
8 Dial Handsfree	Please dial number 031-66 51 81 on the handsfree phone and await a set of questions.	Visual
9 Nav Zoom	Please zoom out to 15 km level on the RTI map (from 3 km default)	Visual
10 Nav Lang	Please change language on the RTI from Swedish to English.	Visual
11 Email Nav	Please read the email (on paper) taped to the RTI display and await questions about it.	Visual
12 Email Radio	Please read the email (on paper) taped to the Radio and await questions about it.	Visual
13 Email Palm	Please read the email on the Palm and await questions about it.	Visual
14 Email TTS	Please listen to the email in the car speakers (text to speech) and await questions about it.	Cognitive
15 Conv Hand Held	Please answer a set of questions on the hand held mobile phone (posed by an experimenter on the phone at a different location).	Cognitive
16 Conv Hands Free	Please answer a set of questions on the hands free mobile phone (posed by an experimenter on the phone at a different location).	Cognitive
17 Conv Pass	Please answer a set of questions (posed by an experimenter in the front passenger seat).	Cognitive
18 Counting	Please count backwards by 7 from 568 (for 30 seconds).	Cognitive

## Results

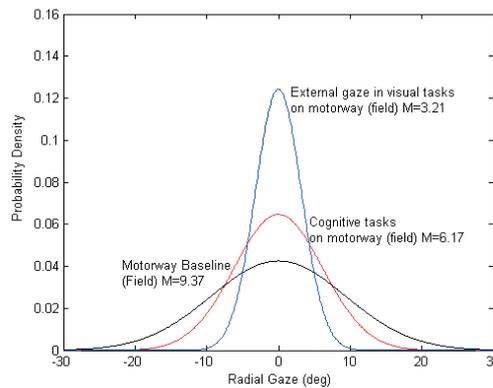
Gaze is directed towards the road center, distant region for a large portion of the time during baseline driving. During visual tasks, a dramatic spatial concentration towards road centre occurs in the returning gaze, and this is accompanied by a sharp reduction in the amount of viewing time on the road centre area. During cognitive tasks, spatial gaze concentration to road centre is also evident but the amount of viewing time on the road centre area is increased.

As can be seen in Figure 5 a) and Figure 6 a), gaze is very concentrated to road center with no fixations on lane markings to the left and right near the vehicle during baseline driving. Figure 5 a) shows that, in addition to the road center area, gaze is spread out on the speedometer and oncoming traffic and signs on the sides of the roads. Figure 1 a) shows that the percent of

viewing time spent on the road center area reduced sharply from an 80% mean (SD = 10%) in baseline driving to a 29% mean (SD = 14%) during visual tasks (PRCtask mean). Gaze to other areas was correspondingly reduced from 20% (SD = 10%) in baseline driving to 7% (SD = 13%) during visual tasks. During cognitive tasks the percentage of gaze time spent on the far road area increased to 88% (SD = 9%), at the expense of other areas which reduced to 12% (SD = 9%). The examples of eye movement data in Figures 6 a) - c) also illustrate these conclusions using examples of gaze in the three situations. The basic difference between a visual task and a cognitive task is whether a visual time sharing behavior between a task object and the road center area is created. This basic difference creates the need for the separate analyses and measures of visual and cognitive tasks as provided below.



a)



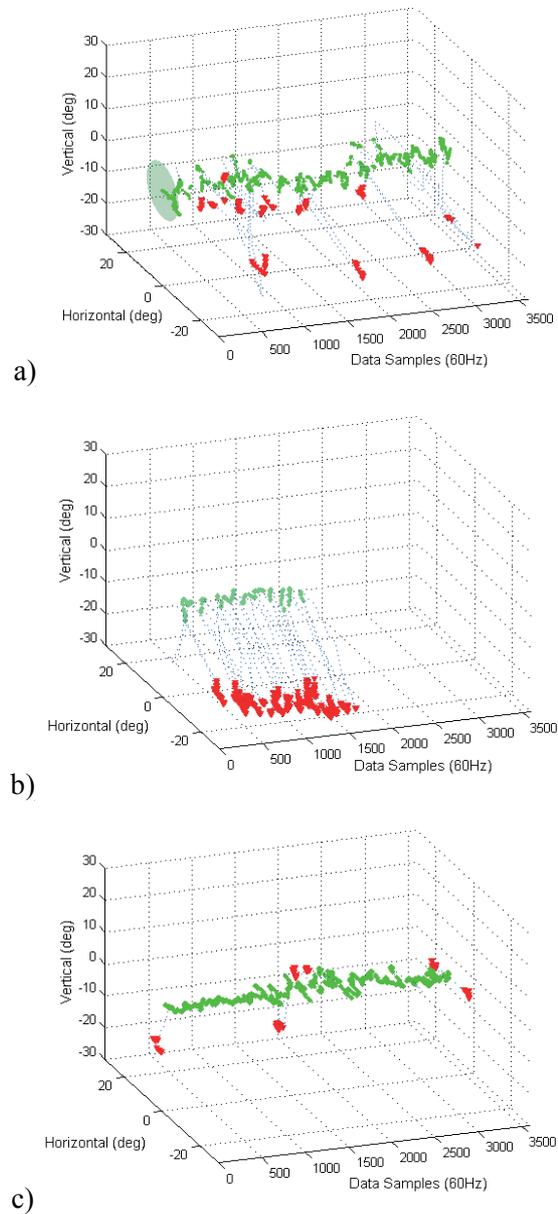
b)

Figure 5 a-b. Spatial Gaze Concentration during baseline driving, visual-, and cognitive tasks as shown in a) mean percent of gaze time on road center (PRCtask), task-related objects, and other areas, and b) mean standard deviations of radial gaze (SDRG in baseline and cognitive tasks, SDREG in visual tasks), plotted as normal distributions.

The spatial variability of gaze, expressed as standard deviation of gaze angles, is also influenced by task type, see Figure 5 b) for results on standard deviation of radial gaze. As expected, gaze became more spatially concentrated to the road center area when performing a cognitive task (e.g.  $M = 6.17$ ,  $SD = 1.62$  for radial gaze) as compared to baseline driving (e.g.  $M = 9.37$ ,  $SD = 1.75$  for radial gaze). However, the concentration effect was even more dramatic in the gaze that was not directed to secondary task related objects during visual tasks ( $M = 3.21$ ,  $SD = 1.26$  in standard deviation of radial external gaze), see Figure 5 b) and 6 b). The reduction in SD of gaze from a baseline driving level, to the cognitive task level, and the further reduction in the visual task level was significant for standard deviation in vertical gaze angles ( $F(2, 28) = 56.20$ ,  $p < .001$ ), horizontal gaze angles ( $F(2, 28) = 66.39$ ,  $p < .001$ ) and radial gaze angles ( $F(2, 28) = 70.51$ ,  $p < .001$ ). All Tukey's HSD pair-wise comparisons between baseline driving, visual tasks, and cognitive tasks were significant ( $p < .05$ ) in all of the vertical, horizontal, and radial gaze angles. So, visual tasks not only caused a greatly shortened road viewing time, they also caused a spatial concentration of the gaze not directed at task-related objects.

### **Visual task results**

Significant main effects when comparing the 13 visual task types were found in all dependent variables (see Table 4 for data and Figure 7 for plots). Respectively, the main effects for Glance frequency, Total glance duration, Total task duration, Mean glance duration, and Number of glance durations exceeding two seconds were  $F(12, 168) = 55.92$ ,  $p < .001$ ,  $F(12, 168) = 67.63$ ,  $p < .001$ ,  $F(12, 168) = 64.42$ ,  $p < .001$ ,  $F(12, 168) = 7.01$ ,  $p < .001$ , and  $F(12, 144) = 4.87$ ,  $p < .001$ . The main effect of Percent road center was significant both when comparing the 13 visual tasks ( $F(12, 168) = 63.18$ ,  $p < .001$ ), and when the baseline driving task was also included in the analysis ( $F(13, 182) = 55.18$ ,  $p < .001$ ). For results regarding post hoc comparisons individual tasks, see study II. As can be inferred in Figure 7, PRC, GF, TGD, and TTD are highly intercorrelated ranging from  $r = .90$  to  $r = .97$ . Note however that PRC can be compared to normal baseline driving.



*Figure 6 a-c.* Examples of gaze scanning behavior during a) a 60 second (3600 data sample) section of baseline driving (the gray circle to the left represents the size of the road center region), b), a visual task, reading an email at radio height in the center console, and c) a cognitive task, listening to an email read out by a computerized text-to-speech voice.

Table 4. Means and Standard Deviations of Dependent Variables for each Visual Task and Baseline Driving.

Task	PRC		PRCtask		SDREG		GF		GF per min		TGD		TGD per min		TTD		MGD		G>2S		
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	
Normal Driving	80,49	10,27																			
1 Climate	77,26	1,18	0,33	0,19	2,82	3,37	2,36	0,87	38,68	8,12	2,17	0,67	37,09	12,28	3,81	1,40	0,98	0,29	0,00	0,00	
2 Engine Temp	76,70	2,03	0,21	0,45	1,95	1,53	3,00	1,60	45,90	16,96	2,96	1,90	41,29	9,44	4,72	3,57	0,95	0,25	0,00	0,00	
3 CD Track	71,83	1,93	0,29	0,11	3,20	2,21	5,18	1,18	31,88	5,44	6,09	1,30	37,70	6,69	9,79	1,69	1,22	0,26	0,09	0,25	
4 Radio Stn	69,67	4,48	0,27	0,13	3,44	3,06	6,38	2,55	31,54	4,72	6,76	2,71	36,26	9,04	12,02	5,05	1,09	0,31	0,31	0,82	
5 Reset Fuel Cons	71,09	4,14	0,34	0,13	4,08	3,91	5,38	1,77	32,47	4,73	4,90	2,59	34,75	8,21	13,45	3,76	0,87	0,19	0,15	0,53	
6 Coin	71,66	5,25	0,44	0,13	2,79	1,14	5,77	2,14	24,69	6,75	6,31	2,89	22,09	7,72	11,62	3,87	1,09	0,29	0,00	0,00	
7 Dial Hand Held	71,28	4,09	0,35	0,13	2,42	1,88	6,54	1,80	30,37	5,52	7,94	1,69	33,06	9,23	11,88	2,36	1,33	0,49	0,31	0,58	
8 Dial Hands Free	69,02	2,01	0,23	0,14	3,71	2,64	6,15	2,39	33,09	6,57	6,92	2,52	41,28	11,68	11,97	5,07	1,15	0,24	0,42	0,59	
9 Nav Zoom	62,95	5,59	0,32	0,09	3,34	2,83	10,71	3,17	30,27	4,67	11,55	3,74	32,95	7,80	21,45	6,36	1,10	0,26	0,50	0,73	
10 Nav Lang	50,39	8,54	0,28	0,12	2,52	1,50	16,08	5,22	27,57	5,11	20,97	5,19	37,01	6,95	34,93	9,51	1,40	0,36	1,75	1,47	
11 Email Nav	53,31	5,20	0,22	0,06	3,96	2,83	13,50	3,09	29,46	5,35	18,15	3,50	39,84	6,39	27,57	4,43	1,42	0,43	1,64	2,09	
12 Email Radio	54,84	5,95	0,25	0,08	3,57	2,04	13,85	3,07	30,62	4,37	16,06	4,42	35,57	7,35	27,14	5,09	1,19	0,33	0,92	1,28	
13 Email Palm	52,44	5,19	0,23	0,10	3,98	2,27	14,29	2,79	29,08	4,20	19,07	3,20	39,30	7,40	30,02	7,74	1,37	0,27	0,64	1,11	

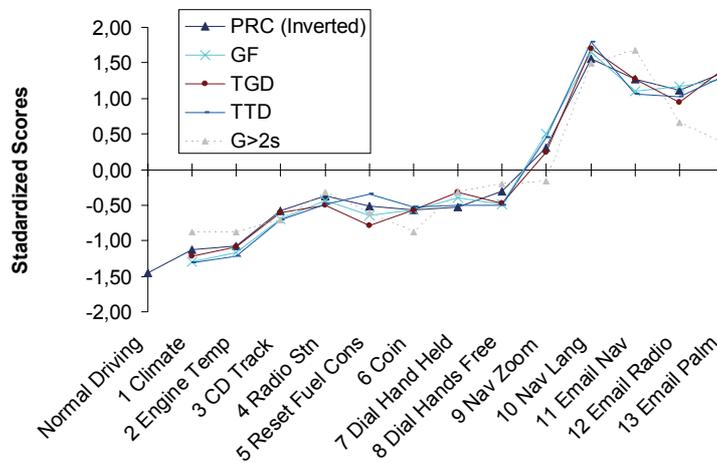


Figure 7. Standardized Scores for Percent Road Center (inverted), Glance Frequency, Total Glance Duration, Total Task Duration, and Number of Glance Durations Exceeding 2 s.

The relationship between task-duration-independent measures and task-duration-inclusive measures was also investigated. The striking similarity among measures that combine task duration effects has been established in correlations and in Figure 7. These strong similarities were also shown to exist between task-duration-independent, rate-based measures in study II. When glance frequency and total glance duration in Table 4 are calculated as rates per time unit (GF per min, TGD per min), they both strongly resemble the effects found in the mean glance duration (MGD) measure. Similarly, when percent road center is not filled out with baseline driving for the rest of one minute, but rather is only calculated for the task duration (PRCtask), it also strongly resembles mean glance duration and the rate based glance measures, especially TGD per minute.

### Cognitive task results

As expected, cognitive tasks created both an increase in gaze time in the road center area and a reduction in gaze variability. Figures 8 a) – f) plot the concentration effect in the five cognitive tasks in comparison with baseline driving. Here we see that the reduction in peripheral areas reduces mostly gaze towards signs, oncoming traffic, and the speedometer.

Significant main effects of the five cognitive task types in comparison with baseline driving were found for all dependent variables (see Table 5 for data). The main effects for Percent road centre, standard deviation of vertical gaze, standard deviation of horizontal gaze, and standard deviation of radial gaze were  $F(5, 70) = 3.15, p < .05$ ,  $F(5, 70) = 4.23, p < .01$ ,  $F(5, 70) = 12.80, p < .001$ , and  $F(5, 70) = 10.81, p < .001$  respectively. For results regarding post hoc comparisons individual tasks, see study II. Horizontal and radial gaze were found to be highly correlated, but radial gaze was slightly higher correlated with PRC than horizontal gaze.

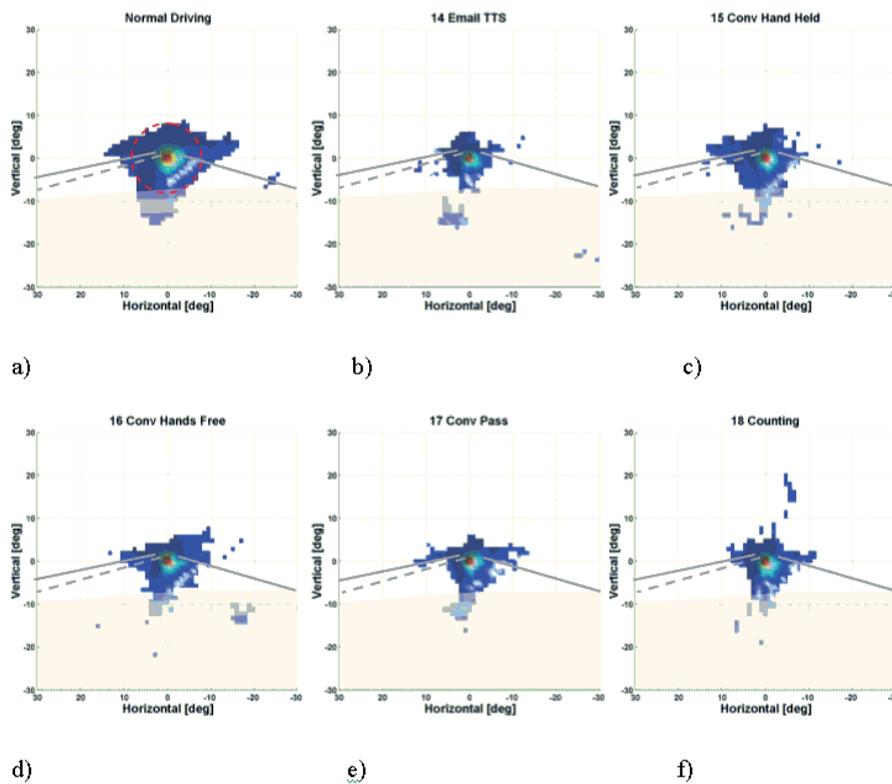


Figure 8 a) to f). Fixation Density Maps of Eye Movements in Baseline Driving and During Cognitive Tasks as Seen from the Drivers Perspective. Frequency in units representing percent of total frequency per bin (one bin is 0.98 square degrees). The plots represent a combination of all the subject data within a particular task. The

superimposed grey area represents the interior of the vehicle and the lines represent road markings. The dashed circle in a) represents the boundary for the road center area.

Table 5. Means and Standard Deviations of all Dependent Variables for each Cognitive Task and Baseline Driving.

Task	PRC		SDVG		SDHG		SDRG	
	M	SD	M	SD	M	SD	M	SD
Normal Driving	80,49	10,27	5,19	0,94	7,75	1,73	9,37	1,75
14 Email TTS	88,86	8,50	4,22	1,90	5,28	2,56	6,84	2,98
15 Conv Hand Held	85,87	9,00	4,64	1,26	5,11	1,70	6,99	1,79
16 Conv Hands Free	86,81	8,12	3,98	0,91	4,73	1,83	6,22	1,91
17 Conv Pass	89,35	7,40	3,49	1,29	3,97	1,56	5,36	1,85
18 Counting	87,94	12,38	3,63	1,93	3,81	1,78	5,42	2,28

## Discussion

The results indicate the importance of the distant road center region for (proactive) visual guidance of driving, and are in agreement with the conclusion that gaze fixation is the main mechanism for trajectory aiming. The results are also consistent with the conclusion that concentration of gaze to the road centre area is the driver's behavioral response to increased demands.

In itself, it is noteworthy that the road center region receives a large amount of viewing time and a large spatial gaze concentration in normal driving. However, the current results also show that this region is increasingly prioritized when both cognitive- and visual tasks are performed. In keeping with literature, the results indicate that drivers spend a dramatically large amount of time with their eyes off road during visual tasks, but the current results also show that the remaining gaze (the gaze not directed towards a task-related object) is highly concentrated to road centre. Visual scanning of the environment is reduced by a dramatic two-thirds in the gaze not directed to an in-vehicle visual task object (SDREG), and is reduced by one third during cognitive tasks (SDRG). Correspondingly, the amount of time spent viewing other areas than the road centre area or a visual task object is reduced by over two thirds in visual tasks and over one third in cognitive tasks. Although gaze also becomes concentrated to road centre during cognitive tasks, there is a crucial difference – the percentage of time spent looking at the road centre area is increased in cognitive tasks and decreased in visual tasks. The conclusion that in-vehicle tasks create a maximization of available gaze towards the road centre region is further verified by similar gaze density plots of visual and cognitive tasks found in study III and Nunes and Recarte (2002).

Spatial gaze concentration has previously only been studied in the context of cognitive, not visual tasks. Recarte and Nunes (2000) and Nunes and Recarte (2002) showed an on-road reduction in standard deviation of radial gaze (calculated from their horizontal and vertical gaze data) from baseline levels between 8.9 degrees and 11 degrees to the largest effect of a cognitive task at 5.9 degrees. Harbluk et al. (2002) found an increase from a baseline of 78.6 percent in a similarly sized road centre area to a cognitive task at 82.7 percent. Tijerina et al. (1995) found that the same arithmetic questions used in tasks 15-18 significantly reduced mirror sampling but did not significantly affect lane exceedences or speed maintenance, that the pattern of steering reversals was not interpretable, and that accelerator holds were only slightly more frequent. Tijerina et al's results fall in line with the general conclusion that cognitive tasks interfere with gaze scanning and reaction time but not with path guidance (Engström et al., 2005; Horrey & Wickens, 2004; and Seppelt & Wickens, 2003).

Regarding the specific effects of individual tasks, the following main conclusions can be made. The simple tasks were not significantly different than baseline driving. There was some evidence that dialing on a hands free telephone is more difficult than a hand held. There was some evidence of longer glances to highly placed displays although the effect was not as large as expected. The medium of conversation was not significant when comparing the hand held phone, hands free phone, and passenger conversations. Rather, the gaze concentration seems to be caused by the cognitive task content itself. Listening to an email being read by computerized text-to-speech and a backward counting task had about the same effect on gaze concentration as the conversations did.

### **Characteristics of eye movement measures and implications for safety**

It was shown that there is an implicit task duration effect present in glance frequency, total glance duration, and percent road centre (calculated with a fixed period). When the task duration effect is removed by converting glance measures to rates and by using PRCtask, they become very similar to mean single glance duration. Importantly, the reason that these rate-based measures are so similar and so flat is because the temporal characteristics of the visual time-sharing behavior between road centre and visual tasks are remarkably constant, irrespective of task length. The consistency implies that drivers typically do not pause during an interaction, but rather continue time sharing until they finish the task. This conclusion is supported by the large proportion of time spent on a display during visual tasks in Figure 5 a). Gaze is time-shared an average of 64% on displays during visual tasks, 29% in the road center area, and 7% on other areas (Figure 5). Task duration is an important factor when evaluating safety because every off-road glance is a safety risk, and longer tasks create more off road glances.

PRC, calculated with the 1 min window, also implicitly weights task duration to a very similar amount as glance frequency and total glance duration, as indicated by very high correlations with glance measures (Table 4 and Figure 3). This time window based approach has also been implemented as a moving time window in real-time distraction recognition algorithms (Victor & Larsson, 2004) and been used successfully to actively counteract distraction by triggering warnings in real time (Victor, 2003; Victor & Larsson, 2004). In a similar approach, Zhang and Smith (2004b) also recommend one minute moving time window based glance measures. They found high correlations between time-based glance measures and driving performance (lane keeping measures, steering entropy, and reaction time measures). Implemented as moving time windows, these time-based measures enable task independent analyses as there is no requirement to define a start or finish for tasks. They can also be used to evaluate interactions with virtually any distraction, even those outside the vehicle.

Collision probability should be expressed in terms of amount of exposure to periods with a certain intensity of driving-information loss. The PRC measure provides exactly such a solution when used as a *moving* time window (see below). Every glance away from road center is associated with reduced path guidance information and increased reaction time to changes in the road ahead. It is especially the simultaneous occurrence of eyes off road, low attention to the road scene, and an unexpected critical event onset that is a crucial factor in reaction time (Green, 2000; Rumar, 1990).

### **The prevalence and significance of the gaze concentration effect**

A more general conclusion of a relationship between increasing situation demands (baseline driving, cognitive tasks, visual tasks) and increasing gaze concentration is made. As shown in study III, gaze concentrates as a function of both incremental increases in task difficulty and incremental increases in steering difficulty in motorway, straight rural, and curved rural sections. The present results also fit well with other research indicating the importance of gazing to the far road center region in driving (e.g. Land & Horwood, 1995; Salvucci & Gray, 2004; study III; Wann & Swapp, 2000; Wilkie & Wann, 2003b), and fit well with general conclusions on the importance of prospective visual information to guide action in everyday motor activities (e.g. Flanagan & Johansson, 2003; Land & Furneaux, 1997; von Hofsten, 2004). Thus, in-vehicle task demand, driving task demand, and driver state factors cause attention to be more focused on the information that is most essential for driving, on prospective heading information at road center. Congruent with Wann and Wilkie's (2004) point attractor model of steering, the far road center region can be seen as an attractor with gaze concentrating to it as situations become more demanding.

As gaze concentrates to road centre as a function of increased situation demands, the driver's ability to scan the environment and perform recogni-

tion and planning tasks is reduced. For example, gaze concentration and performance deterioration in the visual field are shown to occur simultaneously (e.g. Recarte & Nunes, 2003).

The attentional and cognitive demands present in cognitive tasks generally reduce event detection performance but improve- or leave path control performance unaffected (Brookhuis, de Vries, & de Ward, 1991; Engström, et al. 2005; Horrey & Wickens, 2004; Seppelt & Wickens, 2003). Similar to the results from Creem and Proffitt (2001), and other dissociations reported by Goodale and Milner (2004), this dissociation in event detection and tracking performance indicates that cognitive tasks interfere with information provided by vision-for-identification, whereas vision-for-action is more immune. Path control performance during cognitive tasks also benefits from increased gaze concentration to road centre. Gaze simply rests longer on the best information for path control.

In visual tasks, the effect of looking away from the road is added to these attentional and cognitive demands, explaining why the impact on driving performance measures is generally larger from visual than cognitive tasks. In visual tasks, drivers maximize their return glances to road center and reduce their speed to compensate for the effects of visual demand, but these countermeasures are insufficient as lane keeping and other driving performance nevertheless deteriorate (Engström, et al., 2005; study III). However, there is not a complete loss of information during off-road glances, as peripheral vision can be used. Residual information in peripheral vision may partially explain why correlations between lane keeping and off-road glances are high but not perfect.

The present results show that drivers scan less in other peripheral areas during both visual and cognitive tasks, they therefore support the argument that information about the state of the world outside is gradually reduced during secondary tasks (Zwahlen et al., 1988). Unfortunately, a paradox occurs because glancing away from road centre can be supportive of driving, as when looking at signs and potential traffic conflicts that arise in visual periphery as well as detrimental for steering and headway control tasks. In other words, some capability to look away for recognition and planning activities and to monitor the environment should be preserved to ensure safety. If a situation is too demanding to allow the eyes to be taken off road centre, driving behavior should be adapted, for example by slowing down and increasing headway, to make the off-road glances possible and comfortable. Thus, a relationship between safety risk and far road centre viewing time over a longer time scale is proposed (Figure 6 in study II), wherein the lowest safety risk is found when gaze fairly concentrated to road centre but yet still free to roam if need be. Because the dorsal sensorimotor systems have short-lived representations (operate on a short time scale), their representations need to be continually updated. Information acquisition delays should

therefore not be greater than two seconds. Thus there are both short term and long term safety risks.

### **Neural mechanisms behind the gaze concentration effect**

If verbal communication and thought processes are included as competing stimuli to visual stimuli in saliency maps, and if consideration is given to the fact that attentional and eye movement processes are tightly integrated at the neural level, then the reduction in eye movements associated with cognitive tasks is understandable. Attention to inner processes or specific visual stimuli reduce eye movements because these stimuli win competition for attention, in line with Corbetta and Shulman (2002). Stimulus-driven attention mechanisms are suppressed, and salience maps are inhibited in favor for the goal-directed, top-down stimuli (Itti, 2005; Koch, 2004).

The biased competition in attention mechanisms could also be a common cause for the visual field performance deterioration or “tunnel vision” that is associated with increases in attentional demand. It also follows that detection of unexpected events should be more affected than expected events when attention is engaged in something else. This is because unexpected events do not receive goal-directed attention and occur in parts of the visual field where stimulus-driven attention is inhibited, in line with Corbetta and Shulman (2002), Green (2000), Mack and Rock (1998), Rensink (2002a), and Rumar (1990).

An intriguing explanation for gaze concentration effects is that of interference with processing in vision-for-identification and/or the upstream prefrontal cortical regions. Event detection and reactions that involve recognition and planning processes are dependent on conscious, perceptual and cognitive representations of the visual characteristics of objects supplied by vision-for-identification. When cognitive tasks are performed they interfere or compete with the recognition and planning processing in vision-for-identification and/or upstream cortical prefrontal regions that is needed for driving. In contrast to processing based on vision-for-identification information, our vision-for-action visuomotor control systems are largely immune to cognitive interference because they operate on online, principally non-consciously accessible information (Goodale & Westwood, 2004; Koch & Crick, 2001; Rosetti & Pisella, 2002). This explains why tracking behaviour, a predominantly vision-for-action function, is generally not affected by cognitive tasks, and why event detection and reaction time are affected. Indeed, it is dissociations of this kind that lead to the discovery of the two streams (see Milner & Goodale, 1995). Thus, when gaze is concentrated to road centre in visual- and cognitive tasks, it is concentrated on vision-for-action information, with event detection performance and viewing time on other regions being reduced. Because gaze is increased in the road centre region in cognitive tasks, the vision-for-action visuomotor systems are better specified. Although road centre viewing is maximized in the road-refixations in

visual tasks, it apparently is not enough to compensate the performance deterioration caused by off-road glances.

The present findings lend support to the general conclusion that visual guidance of action (vision-for-action) is less susceptible to interference than recognition and planning (vision-for-identification). They also point to a general relevance of prospective information and gaze lead in the visual guidance of continuous actions. Gaze concentration to prospective action guidance information is a behavioral response to increased demands.

## Eye movement effects of different task types, task difficulty, and driving complexity (Study III)

Study III aimed to answer four research questions. How are eye movements influenced by (1) different in-vehicle task types (visual and auditory), (2) increasing in-vehicle task difficulty, and (3) driving task complexity? Lastly, (4) which measures are most suitable and sensitive to these changes in eye movements?

### Design

Eye-movement data were collected by two partners within the EU project HASTE using two different eye-trackers. Data from 119 subjects were collected from four routes: a motorway in real traffic with an instrumented vehicle (Mwy Field), a motorway in a fixed base simulator (Mwy VT Sim), and from rural roads in two different fixed base simulators, one in Sweden (Rur VT Sim) and one in Canada (Rur TC Sim). The rural routes were additionally analyzed with respect to whether the road was straight (Rur VT Sim Straight and Rur TC Sim Straight) or curved (Rur VT Sim Curve and Rur TC Sim Curve) according to road complexity level.

Increasingly demanding in-vehicle tasks by means of artificial, or surrogate In-vehicle Information Systems (S-IVIS) were used. The visual S-IVIS task, referred to as the visual task, was designed and pre-tested to produce three incremental levels of difficulty (SLv1-3). Briefly described, the subjects were presented with matrices of arrows on a screen positioned to the right of the steering wheel. Each task lasted 30 seconds and consisted 6 presentations of matrices, one every 5 seconds, for each of the three difficulty levels. The auditory task was an auditory working memory task, called auditory continuous memory task (aCMT), wherein the driver was required to remember 2, 3 or 4 target sounds, corresponding to SLv1, 2, and 3. After being presented the target sounds, a series of 15 sounds are played back for the driver to keep track of the number of times each target sound was played

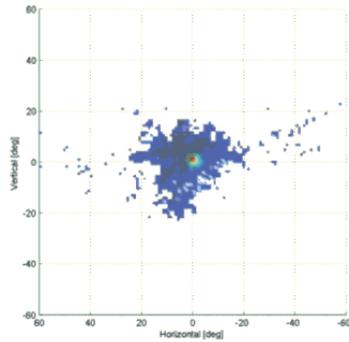
over a span of 30 seconds. As all tasks were 30 seconds, the effect of task length was controlled for.

Seven measures were analyzed: Mean glance duration, Percent glance durations exceeding 2 seconds, Standard deviation of glance duration, Glance frequency, Total glance duration, Percent road centre, and Standard deviation of gaze, see Table 2 for definitions. PRC was calculated for a fixed time of 30 seconds, and thus it is equivalent to both the PRC and PRCTask measures in study II. In the data analyzed at Volvo (the Mwy Field and VT Sim experiment data) the road centre area was defined as a circle of 16 degrees diameter, centred around the road centre point determined as the mode. In the data analyzed at University of Calgary (Rur TC Sim data), the road centre area was defined as a 20° (horizontal) x 15° (vertical) rectangular area centered around the road centre point, determined as subject's mean fixation. Mean fixation was calculated by including only those fixations that were in the forward view (i.e. the central projection screen). For the TC Sim data set standard deviation of gaze was calculated using gaze position instead of gaze angle.

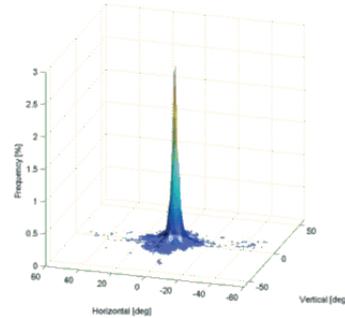
Univariate ANOVAs, with Subject included as a random factor, were used to test the statistical significance at a 5% level in the dependent variables. The main independent factors investigated were S-IVIS task complexity (BL, SLv1, SLv2, and SLv3) and road complexity (the straight and curved road sections on the rural road). Sidak adjustments were used for post hoc pairwise comparisons of means. Differences between test environments (the two different static simulators and the field), between S-IVIS types (visual vs. auditory), and between motorway and rural road types were analyzed qualitatively only.

## Results

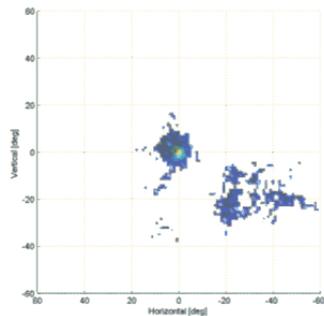
The first striking feature about the gaze data, as can be seen in Figure 9 a) to d), is that it is very concentrated to a region straight ahead. Figure 9 b) shows a side view with this concentration apparent in normal baseline driving on a motorway in the field. The two baselines, Figures 9 a) and b), are collected in different sections of the same motorway and are very similar and are very similar to that in Figure 8 a). As can be seen in Figure 9 c), this general pattern changes quite dramatically when a visual task is introduced, as it did in Figures 5 a) and 6 b). Thus, a concentration effect is seen in the upper on-road cluster and clusters of fixations are introduced where the in-vehicle visual task display is. During the auditory task in Figure 9 d), a concentration effect is also visible and quite similar to Figures 8 b) – f).



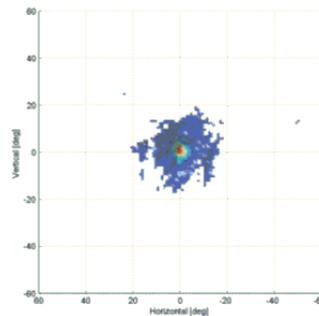
a) Visual task baseline Mwy Field



b) Auditory task baseline Mwy Field



c) Visual task SLv3 Mwy Field



d) Auditory task SLv3 Mwy Field

Figure 9 a) to d). Examples of fixation density plots of eye movements in different conditions as seen from the driver's perspective. The plots represent a combination of all the different subject data within a particular condition. Frequency is normalized to units representing percent of total frequency.

Pairwise Sidak post-hoc comparisons for the visual task in Table 6 show that all measures were sensitive to visual task difficulty (for main effects, see study III). Percent road centre was sensitive in 22 (92%) of 24 pairwise comparisons, including comparisons with baseline. Total glance duration was sensitive in 7 (78%) of 9 comparisons, followed by Mean glance durations at 6 (67%) of 9, and Glance frequency at 5 (56%) of 9. Standard deviation of glance duration and Percent of glances exceeding 2 seconds were both sensitive in 4 (80%) of 5 comparisons. The pairwise comparisons for the auditory task in Table 6 show that both Standard Deviation of Gaze and Percent Road Center were sensitive to task difficulty. Standard deviation of gaze was most sensitive in 12 (50%) of 24 pairwise comparisons and Percent road centre was sensitive to 4 (17%) of 24 comparisons.

Table 6. Post-hoc pairwise comparisons for all dependent measures.

	SLv1				SLv2				SLv3			
	Mwy VT Sim	Mwy Field	Rur TC Sim	Rur VT Sim	Mwy VT Sim	Mwy Field	Rur TC Sim	Rur VT Sim	Mwy VT Sim	Mwy Field	Rur TC Sim	Rur VT Sim
BL												
Percent Road Centre (visual task)	*	*	*	*	*	*	*	*	*	*	*	*
Percent Road Centre (auditory task)	*	*	-	-	*	-	-	-	*	-	-	-
Standard Deviation of Gaze	*	*	*	*	*	*	*	*	*	*	*	*
SLv1												
Mean Glance Duration					*	-	n/a	-	*	*	n/a	*
Standard Deviation of Glance Duration					*	n/a	n/a	-	*	n/a	n/a	*
Percent Glances Exceeding 2 seconds					*	n/a	n/a	-	*	n/a	n/a	*
Glance Frequency					*	-	n/a	*	*	-	n/a	*
Total Glance Duration					*	-	n/a	*	*	*	n/a	*
Percent Road Centre (visual task)					*	-	*	*	*	*	*	*
Percent Road Centre (auditory task)					-	-	-	-	-	-	-	-
Standard Deviation of Gaze					-	-	-	-	-	-	-	-
SLv2												
Mean Glance Duration									-	*	n/a	*
Standard Deviation of Glance Duration									-	n/a	n/a	*
Percent Glances Exceeding 2 seconds									-	n/a	n/a	*
Glance Frequency									-	*	n/a	-
Total Glance Duration									-	*	n/a	*
Percent Road Centre (visual task)									-	*	*	*
Percent Road Centre (auditory task)									-	-	-	-
Standard Deviation of Gaze									-	-	-	-

'\*' =  $p < .05$ , '-' =  $p > .05$ , 'n/a' = not applicable

Focus here is on the Percent road centre and Standard deviation of gaze measures, please refer to Study III for results regarding other measures. Percent road centre decreased sharply from baseline when a visual task was performed and decreased further as task difficulty increased, as can be seen in Figure 10. The drop from baseline driving to SLv1 was comparatively the largest. Significant task difficulty main effects were found in all data sets; Mwy VT Sim  $F(2, 38) = 15.48, p < .001$ , Mwy Field  $F(3, 49) = 90.27, p < .001$ , Rur TC Sim  $F(3, 66) = 68.71, p < .001$ , and Rur VT Sim  $F(2, 41) = 43.41, p < .001$ . As shown in Table 6, subjects looked significantly less at the road centre area in all visual tasks (SLv1-3) as compared to baseline

driving. Additionally, subjects looked significantly less at the road centre in SLv3 than in SLv1 in all data sets. SLv2 produced significantly less gazing at the road centre than SLv1 in all data sets except Mwy Field. Finally, SLv3 caused significantly less gazing at road centre than SLv2 in all data sets except Mwy VT Sim. Thus, PRC was able to discriminate between each difficulty level as well as baseline. Subjects looked significantly more at the road centre area in curves than in straight sections in both the Rur VT Sim  $F(1, 52) = 5.23, p < .05$  and Rur TC Sim  $F(1, 25) = 5.42, p < .05$  data sets.

Percent road centre increased when the auditory tasks were performed, as can be seen in Figure 10. However, task difficulty main effects were only found in the Mwy VT Sim  $F(3, 65) = 82.10, p < .001$ . As shown in Table 6, subjects looked significantly more at road centre in SLv1 than in baseline driving in the Mwy VT Sim and Mwy Field data sets. SLv2 and SLv3 also produced significantly more gazing at road center than baseline in the Mwy VT Sim data set. Subjects looked significantly more at road center in curves than in straight sections in the Rur TC Sim data set  $F(1, 23) = 6.34, p < .05$ , the same finding as PRC in the visual task. The Rur VT Sim data showed the same trend. The baseline PRC is highest in rural curves, followed by rural straight sections, the simulated motorway and finally the field motorway. This same trend as is evident in the visual PRC data.

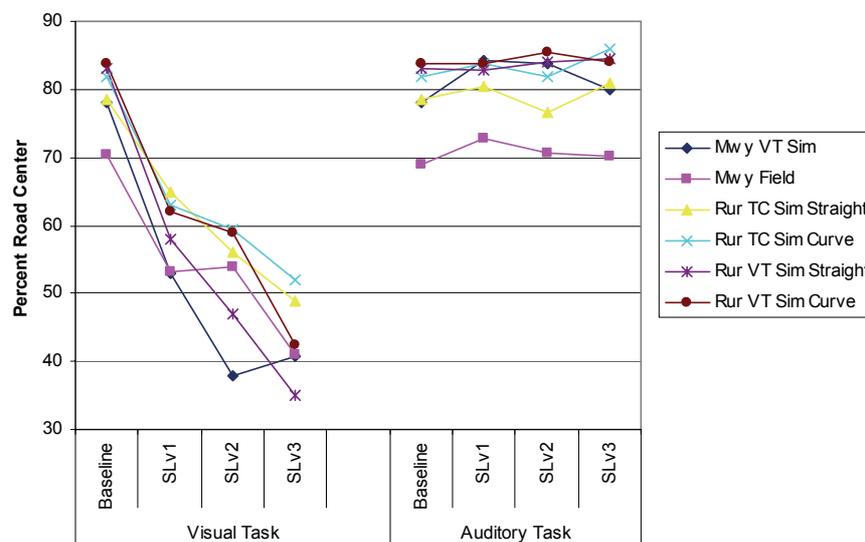


Figure 10. Percent road centre in the visual and auditory tasks, including baselines and the three task difficulty levels (SLv1-3) per task type.

Standard deviation of gaze was reduced when the auditory tasks were performed, as can be seen in Figure 11. Significant main effects for task difficulty were found in all data sets, Mwy VT Sim  $F(3, 69) = 8.34, p < .001$ ,

Mwy Field  $F(3, 65) = 7.61, p < .001$ , Rur TC Sim  $F(3, 70) = 10.86, p < .001$ , and Rur VT Sim  $F(3, 65) = 3.88, p < .05$ . The reductions from baseline ranged at most between 23 and 37%. Standard deviation of gaze was significantly reduced in all auditory tasks compared to baseline, as can be seen in Table 6. There was a tendency that gaze was more concentrated in curves than in straight sections. Also in line with the other results, gaze is more concentrated in rural driving than in motorway driving. The difference between the field motorway and the simulated motorway is also very evident. All in all, the gaze concentration effect was more pronounced in the standard deviation of gaze measure than in the PRC measure.

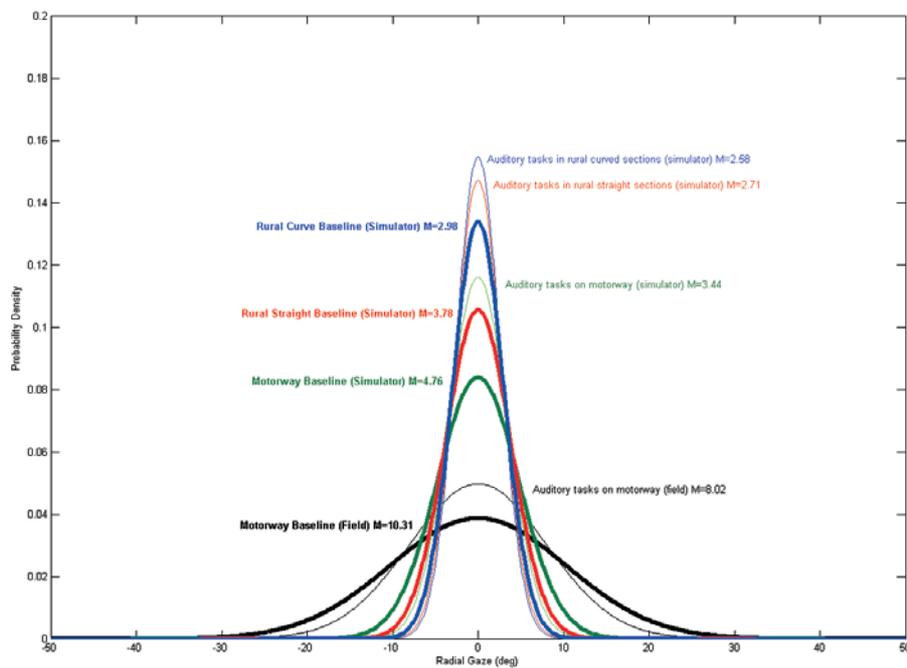


Figure 11. Standard deviation of gaze in baselines and auditory tasks (one mean of three task difficulty levels per baseline). Standard deviation of gaze values are plotted as normal curves.

## Discussion

### Different in-vehicle task types

The visual and auditory task types caused very different types of changes to eye movement behavior. Visual tasks cause drivers to look away from the road, and auditory tasks cause drivers to concentrate their gaze to the road center area. The baseline and auditory task density plots in Figure 9 are very similar to those plotted in Nunes and Recarte (2002) and Figure 8, showing a concentration of gaze in auditory tasks. Figure 9 also shows the same gaze concentration of on-road glances in visual tasks as was found in study II,

Figure 5 b). One important similarity between the visual and auditory task is that the gaze concentration effect intensifies quite dramatically when a visual task is introduced. So, in addition to losing information about the driving environment when looking at an in-vehicle display, the returning glances to the road are aimed increasingly at the road centre during a visual task. Although no metrics of the spatial concentration of gaze in the forward view during the visual task are provided here, the effect is evident in Figure 9 c) and was significant in study II.

### **Increasing in-vehicle task difficulty**

Consistent with expectations, the data shows that increases in task difficulty produces both an increase in display viewing time and a gaze concentration to the road centre area in remaining gaze. Importantly, Study III shows that Percent road center and glance measures increase with task difficulty *even when task duration is held constant*.

As the visual task became more difficult, all measures clearly show that drivers look more at the display and less at the road ahead. Although this is a rather intuitive finding and is supported by all literature, the task difficulty is systematically controlled here. Study III also shows that Percent Road Centre is most sensitive to task difficulty, followed by Total Glance Duration.

Although the results show significant gaze concentration caused by the auditory task (measured by SDG and PRC) in comparison with baseline, they do not show significant increases in gaze concentration *between* auditory tasks (SLv1-3). Significant differences between different auditory or cognitive tasks in gaze concentration have not been shown previously in other research either, only differences with baseline as in study II. In general, the changes to eye movements caused by the auditory task were weaker in magnitude than the effects of the visual task. In this particular task, one explanation for the apparent reduction in more difficult tasks as compared to the easier task (SLv1) could be that drivers may have found the task too difficult and stopped trying on some of the sounds, thereby freeing up cognitive resources, which in turn counteracted the gaze concentration effect.

### **Driving task complexity**

Drivers adapt their eye movement behaviour to complexity in the driving environment, see Figure 11. In general, they increase viewing time in the road centre area when driving task difficulty increases, as evidenced by the spatial gaze concentration being highest in the rural curves, followed by rural straight sections, the simulated motorway, and the field motorway. In curves, drivers consistently show shorter single glance durations, fewer glances exceeding 2 seconds, less variation of single glance durations, less total glance duration, and higher proportion of fixations on the road centre region. These findings are in line with previous research (e.g. Wierwille, 1993a, 1993b; Senders et al., 1967).

The field motorway had a lower baseline PRC than the other road types, indicating that the path control task did not require as much visual guidance and/or that there was simply more to look at. Notably, drivers were also less susceptible to take their eyes off the motorway in the field to do the in-vehicle tasks than they were in the simulated roads, as indicated by the disappearance of check glances, shorter glance durations and less frequent glances to the display. Perhaps this reflects a difference in perceived risk between these settings. More research is needed.

### **Sensitivity and suitability of eye-movement measures to changes in eye movements**

In general, eye-movement measures were found to be highly sensitive to the demands of visual and auditory in-vehicle tasks as well as driving task demands. However, Percent road centre and Standard deviation of gaze, were found to be more sensitive, more robust, more reliable, and easier to calculate than established glance-based measures.

Percent road centre was the most sensitive measure and was alone in being able to compare both visual and auditory task types with baseline driving. The most sensitive visual task measures were those where glance duration and frequency are implicitly combined - Percent road centre and Total glance duration. But if changes in glance strategy are of interest then Glance frequency and Percent glance durations exceeding 2 seconds or Mean glance duration are needed. In the auditory task, Standard deviation of gaze was more than twice as sensitive as Percent road centre. PRC only picked up the larger effects in the motorway data sets. Similarly, Harbluk et al. (2002) and study II both found effects on the motorway. SDG is better because it is sensitive to the smaller changes in gaze concentration in the rural environment. Yet curiously, it was in PRC that differences between curved and straight road types were picked up. In the HASTE final report (Carsten et al., 2005), PRC was selected as one of the six recommended most sensitive measures to detect the effects of a particular system on driving (the others were subjective ratings, mean speed, high frequency steering, minimum headway, and Peripheral Detection task reaction time).

Glance based measures have been notoriously difficult and/or time consuming to collect and analyze. This is a main reason why they are not as frequently used as vehicle performance measures. Measurements of glances are sensitive to signal noise because a glance is a rather long, cohesive, ordered sequence of data. The gaze signal noise in eye-trackers increases at larger visual angles, as the eye rotates away from the camera. Therefore, sophisticated analysis procedures are required so that the noise does not break up one glance into several smaller glances. Importantly, measurement accuracy is better in the forward view, and therefore, it makes more sense to measure where the accuracy is best (as in PRC) rather than at the limits of measurement (measuring glances to displays at large visual angles). Other

factors, such as seating position and movement front- or backwards greatly change the visual angles to eccentric displays, whereas the central area remains central. In contrast to glance-based measures, PRC and SDG do not rely on measuring an ordered sequence of data, only on measuring single data points, and thus they are more robust and easier to measure. Note that PRC and SDG, could have been based on raw, unfiltered, unsegmented gaze data. Also, the fact that the PRC and SDG measures were successfully implemented on different eye trackers using different analysis software leads to the conclusion that the measure is robust.

#### **A note on eye-tracker accuracy**

A potential concern with the FaceLab eyetracker system used in studies I, II, and III regards its accuracy. A large amount of effort was put into developing an analysis procedure that would ensure good data quality when collecting data from drivers in naturalistic environments. Additionally, face markers were used to improve eye tracker performance. Study I, Larsson (2003), and Victor and Larsson (2004) bear witness to these efforts in validating and developing data quality management algorithms. Additionally, FaceLab data was successfully compared with an ASL eyetracker in study II, and several different analysis software implementations found similar results in study III, and in later projects (e.g. Johansson et al. 2005). If poor data or noisy data were obtained, there would have been larger variation in Figures 8 and 9 as compared to Nunes and Recarte (2002), the SD of gaze results would not have been as similar as what was previously reported, Recarte and Nunes (1999, 2003), and the PRC data would not have been as similar as found in comparison to Harbluk, et al. (2002).

#### **Summary**

In-vehicle tasks requiring vision cause drivers to look less at the road ahead and look more often, for longer periods, and for more varied durations at the in-vehicle display. Auditory tasks cause drivers to look more concentrated at the road center area at the expense of glances to the road scene periphery, e.g. signs, and at the expense of glances inside the vehicle, e.g. the speedometer. Density plots of eye movement behavior also show a dramatically increased concentration of gaze to the road center area when the drivers look back to the road scene while performing visual tasks. The results also clearly show that drivers adapt their eye movement behaviour to the driving task complexity, as shown in differences between the rural curves and straight sections, rural and motorway road types, simulator and field motorways, and different simulators. In general, drivers increase viewing time in the road centre area when demands increase; either as a consequence of looking away from the road during visual tasks, as a consequence of cognitive or auditory task demands, or as a consequence of increased driving task demands.

# General Discussion

## Summary of main empirical contributions

The following main conclusions can be made from the empirical work:

1. Gaze is focused increasingly at road centre as a function of increased situation demands. This gaze concentration represents the driver's behavioral response to improve trajectory- and collision control performance by increasing attention to high quality vision-for-action information found at road centre.
2. Three types of situation demands were shown to affect gaze concentration:
  - a. *Visual tasks*: When drivers look back at the road when performing visual tasks, gaze is very highly concentrated to road centre.
  - b. *Cognitive tasks*: While cognitive and auditory tasks are performed, gaze becomes concentrated to road centre.
  - c. *Driving task*: Gaze concentrates to road centre as the driving task becomes more difficult, as shown when comparing motorway, rural straight, and rural curved sections.
3. A direct consequence of gaze concentration is that the amount of time spent viewing other areas (not road-centre and not a task-related object) is reduced by over two-thirds in visual tasks and over one-third in cognitive tasks. This supports the argument that information about the state of the world outside is gradually reduced during secondary tasks (Zwahlen et al., 1988).
4. Specific task differences:
  - a. Reading emails and SMS messages, and changing zoom level and language settings on a navigation system are significantly different than common in-vehicle tasks such as changing radio stations or dialing a number on a mobile phone and are judged unsafe.
  - b. There was some evidence that dialing on a hands free telephone is more difficult than a dialing on a hand held telephone.
  - c. There was some evidence of longer glances to highly placed displays although the effect was not as large as expected.

- d. Talking on a hand held phone, a hands free phone, and with a passenger were not significantly different from one another.
  - e. Similar effects were found when comparing listening to an email being read by computerized text-to-speech, a backward counting task and the hand held-, hands free-, and passenger conversations.
5. The Percent Road Centre (PRC) measure was developed here and was shown to be the best performing eye-movement measure for visual tasks. It was more sensitive, more robust, more reliable, and easier to calculate than established glance-based measures. For cognitive and auditory tasks, Standard deviation of gaze and PRC were the most suitable measures.
  6. Both the glance-based- and the gaze concentration measures indicated that visual tasks cause drivers to look increasingly more often, for longer periods, and for more varied durations away from road centre as task difficulty increases, even when task duration is held constant.
  7. An automatic eye-movement analysis method was developed and successfully validated in comparison with a manual video transcription method.

## Performance dissociations in path control and reaction time measures are associated with eye movements

It is important to consider the effects on measures of action performance in driving that correspond with measures of eye movements. Using driving performance data collected in the same Mwy Field and VT Sim studies that are reported in study III, Engström et al. (2005) found that the amount of time spent looking at road centre has a strong relationship to path-control measures. As expected, when drivers look away from the road in the visual task, lane keeping performance deteriorates. But Engström et al. also show that the increased spatial gaze concentration in the auditory tasks is associated with *improved* lane keeping performance (reduced lateral variation and an increased number of steering micro corrections). That is, the mere effect of the eyes spending more time on road centre improves lane keeping performance, despite increased cognitive load from an auditory task. Similar results showing improved lane keeping performance when performing the cognitive task in used study III were consistent across six sites in simulators and field settings using the same standardized HASTE experimental design (see Östlund et al. 2004 for details).

This finding is in line with other research (Brookhuis, de Vries, & de Ward, 1991; Horrey & Wickens, 2004; Johansson, et al. 2005; Seppelt & Wickens, 2003; and Tijerina, et al. 1995). Similarly, Recarte and Nunes (2000, 2003) found that the gaze concentration effect, when caused by cognitive tasks, is associated with general reaction time interference in event detection capability.

One further piece of evidence showing a similar dissociation between path control and reaction time measures is reported in Victor and Åberg (2005), see Figure 12. Victor and Åberg asked participants to drive a simulator while continuously fixating a point (no foveal task was present) inside the vehicle for 20 seconds. Full covert (peripheral) attention was devoted to staying in lane and to reacting with the brake to the sudden appearance of a large red square the size of a small car at 30m ahead. The inter-stimulus random time variation of the red square ranged between three and five seconds and therefore was a large, salient, highly expected event. Six fixation positions at increasingly peripheral visual angles were tested during three repetitions. During the first two repetitions the driver simply fixated a peripheral point, but during the last repetition, drivers additionally performed a cognitive task wherein they repeatedly subtracted by seven as quickly as possible from a three digit starting number. As can be seen in Figure 12 a), lane keeping performance (the amount of time outside of lane) improved significantly in all three repetitions, despite the addition of the cognitive task. However, although a similar improvement trend was found in brake reaction times between the first and second repetitions, this trend was significantly *reversed* by the cognitive task in the third repetition. That is, reaction time deteriorated to a poorer level than was found during the first repetition because of the addition of the cognitive task, despite there being an improvement between the first and second repetitions, see Figure 12 b). There was general interference in RT independent of visual angle, including at fovea.

Figure 12 a) also shows the effect of visual angle on lane keeping performance. Even though drivers were able to commit their full attention to lane keeping while looking away, performance was sub-optimal and deteriorated with visual angle, but they were able to adapt by learning perhaps to put more weight on peripheral cues such as alignment of the windshield with road edges. The brake reaction stimulus Figure 12 b) showed small effects of visual angle because the stimulus was large and involved a large discrete visual transient change at onset, a smaller stimulus would have shown a larger effect of visual angle eccentricity.

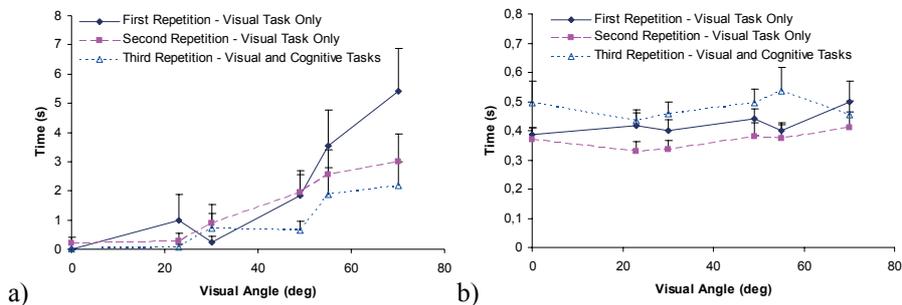


Figure 12 a) – b). Results showing a dissociation between a) lane exceedency, a lane keeping measure, and b) brake reaction time. Lane exceedency is the number of seconds with at least one wheel outside of the lane during a 20 second trial. Brake reaction time is the reaction to the appearance of a large red square the size of a small car at 30m ahead. Error bars indicate standard error. The inter-stimulus random time variation ranged between three and five seconds and was expected. Visual angle represents the angle between road and stimulus centre (0 deg) and a fixation point inside the car.

When the eye movement results are considered together with associated vehicle performance findings the general conclusion is that visual guidance of path-control (vision-for-action) is not as susceptible to interference from cognitive or auditory tasks as recognition and planning tasks (vision-for-identification). Note that planning-related eye movements, as well as recognition of many events and hazards, generally require eye movements away from road centre region towards peripheral objects (e.g. signs, other vehicles, pedestrians). As shown in both study II and III, it is precisely the off-road centre eye movements that are reduced by cognitive and auditory tasks. Thus, both a) reaction time tasks involving peripheral detection (covert attention) and b) recognition and planning tasks requiring eye movements (overt attention) are negatively affected by cognitive and auditory tasks. Driving performance results (Engström et al., 2005; Östlund, et al., 2004; Victor and Åberg, 2005) clearly show that there is a dissociation between lane keeping and reaction time tasks with regard to the effects of a cognitive or auditory task.

## Main contributing factors to inattention in crashes

The following general conclusions can be made regarding important factors influencing inattention-related crashes, see also Figure 1.

### The stimulus saliency factor

Detectability is highly dependent on objective stimulus saliency properties such as size, color, contrast, movement, and luminance (Neumann & Sand-

ers, 1996; Stelmach, Bourassa, & di Lollo, 1984). This is a rather obvious conclusion, but nevertheless important to remember.

### **The shutter factor**

Blinks, saccades, and temporary occlusions are periods of vision loss that mask visual transients and impair event detection (Rensink, 2002) in stimulus-driven attention (Corbetta & Shulman, 2002). We do not have access to a continuous stream of visual information. It is possible that gaze concentration, longer fixations, and blink suppression may have to do with the brain optimizing the visual system so that the “interpolation” associated with saccades and blinks is minimized.

### **The eccentricity factor – looking away**

The eccentricity factor is the effect of stimuli falling on visual periphery instead of at fixation. This factor is of particular concern with in-vehicle visual tasks because drivers spend a dramatically large amount of time with their eyes off road during visual tasks as shown in studies I, II, and III.

The fixation mechanism is the main mechanism for trajectory aiming, and fixating a point on the future path gives the best coordinates (e.g. Wann & Swapp, 2000). Looking away from the road gives poor quality information in peripheral vision causing reduced driving performance in path keeping and reaction time measures. Looking away depletes the short-lived, spatially accurate information resulting in an increased need to collect spatially accurate information in the road-returning glances during visual tasks. This prioritization of road centre information in returning glances is a good example of action-driven attention as outlined above – the action-control of path and headway (basic vision-for-action tasks) *drives* the need to look and attend to road centre. A by-product of this prioritization of the distant region is that scanning of the environment is reduced, and therefore also information regarding the state of the world outside.

The visual eccentricity effect is more of a structural limitation in the sense that we only have one visual resource or one fovea rather than a central processing limitation. This kind of sensorimotor information depletion is a “mechanical” or anatomical effect.

### **Cognitive factors**

In natural driving, and visual-, auditory-, and cognitive tasks there is an added effect of many cognitive factors. I will not attempt to summarize them all. However, important factors identified here include the functional difference between vision-for-action and vision-for-identification, attention, prediction, and cognitive demands from increased task difficulty.

### *The two functions of vision*

Perhaps the most important theoretical contribution of this thesis comes from applying the two-stream vision framework to eye movements and driving performance. Eye movements are motivated by the need to improve acuity and cortical processing power to a) identify objects and events, and b) guide actions (Findlay & Gilchrist, 2003; Land & Hayhoe, 2001; Leigh & Zee, 1999; Milner & Goodale, 1995). Thus, the two visual functions produce conflicting demands for use of the fovea, see Figure 13.

On the one hand, fixations are used proactively to gather the preparatory information needed to guide actions. Here, fixations are gaze anchor points that provide metrically precise information to visuomotor systems to give them true real-world position of a target relative to body. Information is preferably collected from the distant region with the fovea, and from the near road region with peripheral vision. Thus, vision-for-action uses information from across the retina (Baizer et al., 1991; Goodale & Humphrey, 1998), see Figure 13.

The fundamental importance of fixations for action guidance is demonstrated by the way gaze concentrates to the distant path region as demands increase from the driving task, and visual- and cognitive tasks. Gaze concentration in road-returning glances during visual tasks occurs because action-driven attention is serving the needs of vision-for-action. During cognitive and visual tasks, visual guidance of path-control is not as susceptible to interference both because it utilizes more direct perception-action links, but also because gaze spends more time on the distant path region.

On the other hand, fixations are used to inspect and identify stimuli serving the needs of higher cognitive functions. Identification of objects and events outside the vehicle items (e.g. signs, roadside features), and identification of in-vehicle content (e.g. email) need to use the fovea because of the need for acuity but also because receptive fields in vision-for-identification are centered around the fovea (Baizer et al., 1991; Goodale & Humphrey, 1998).

The dissociation that occurs in lane keeping- and brake reaction time performance during cognitive and auditory tasks corresponds to these functional differences in vision. Interference with semantic processing in vision-for-identification affects reaction time during cognitive tasks, probably within goal-setting. These interpretations gain support from similar dissociations between ventral and dorsal streams found in other domains (e.g. Creem and Proffitt, 2001; Milner and Goodale, 1995).

A caveat is in order however. The two functions of vision are interconnected and only semi-independent. Fixations therefore involve both types of processing. A continuum between a pure sensorimotor fixations and pure cognitive identification fixations is suggested.

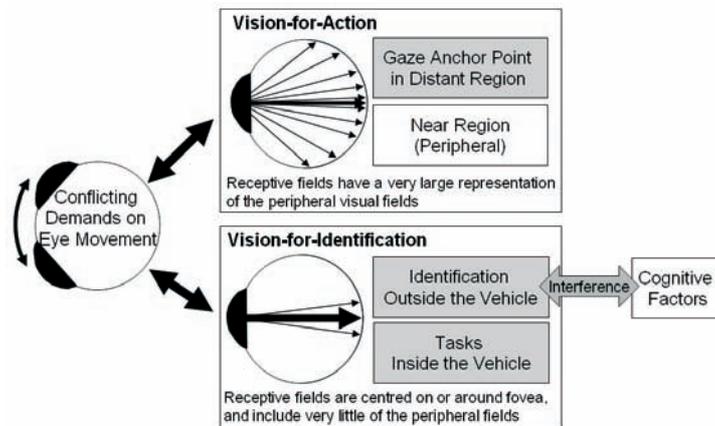


Figure 13. Conflicting demands on eye movement while driving and performing secondary tasks. There is a fundamental conflict or incompatibility between the use of foveal vision for action control (path- and headway control) and the use of the fovea for other tasks, such as looking at visual objects inside the vehicle (e.g. reading an email) and outside the vehicle (e.g. roadside objects) which causes visual time-sharing. Cognitive and auditory tasks interfere with the use of the fovea for identification tasks and reduce the amount of fixations outside the distant region. As difficulty increases (e.g. in the driving task, in visual tasks, or in cognitive tasks) the conflict becomes larger. Grey boxes represent foveal vision.

#### Attention and expectation

The interaction between action-driven-, stimulus-driven-, and goal-directed forms of attentional selection is another important cognitive factor. Goal-directed attention has a large impact on performance when increased difficulty of visual tasks causes more goal-directed attention to be paid. Another explanation for the reduction in scanning eye movements is that attention to thought processes wins competition for attention over peripheral visual stimuli, thereby reducing eye movements. Stimulus-driven attention mechanisms are suppressed, and salience maps are inhibited in favor for goal-directed attention to thoughts (Itti, 2005; Koch, 2004). This could be a combinatorial effect to the interference in vision-for-identification processes as described above. More research is needed.

Expectancy and similarity to currently attended items are also identified as having a large impact on stimulus detectability (e.g. Mack & Rock, 1998; Most et al., 2005), leading to the conclusion that detection of unexpected events is more affected by engagement of attention than detection of expected events. Basic attention research (e.g. Corbetta and Shulman, 2002) and accident research (e.g. Green, 2000; Neale et al., 2005) seem to be in agreement on conclusions regarding expectation.

It is when stimulus saliency-, shutter-, visual eccentricity-, and cognitive factors combine that safety is most compromised, for example when there is a simultaneous occurrence of eyes off road, poor attention to the road scene, and an unexpected critical event.

In closing, it is worth noting that it is very difficult to be specific in driving research without missing important contributing factors. Many factors that were not considered here undoubtedly play an important role. One particular factor not considered is the role of biomechanical interference from manual interactions. Also individual differences (e.g. skill, age, style, etc), impairment (e.g. drowsiness, alcohol), and many higher-level aspects of decision making in driving play important roles. Likewise it is difficult to be all-encompassing yet specific. It will have to be a task for the future to set this work into more general frameworks.

## Attention Support – How to keep eye and mind on the road in practice

How do you recognize when a driver is inattentive? The driver could be looking at the wrong place, blankly staring straight ahead, or not looking at all. Perhaps sudden braking or a quick swerve alerts you to the problem. What would you do as a passenger if you realized your driver is inattentive? Shout and point?

In closing, it will be shown how the knowledge presented here regarding vision, eye movements, attention, and prediction can be applied to in-vehicle products to reduce the number of crashes caused by inattention. In Sweden, universities are charged with a "third task" of encouraging the impact of research results in society. The results outlined in this section are such an example.

It is proposed here that a reduction in the effects of inattention can be achieved by giving vehicles the ability to recognize driver inattention and the means to help drivers to stay attentive. A number of functions will be presented that are intended to work in real-time in future vehicles, and are described in greater detail in (Arensberg, 2004; Claesson, 2003; Engström & Victor, 2005; Larsson & Victor, 2005; Victor, 2003, 2004; Victor & Jarlen-grip, 2005a, 2005b; Victor & Larsson, 2004, 2005; Victor 2000a, 2000b, 2001a, 2001b).

### Distraction alerts

A real-time algorithm was developed to identify visual and cognitive distraction in real-time (Larsson, 2003; Larsson & Victor, 2005; Victor and Larsson, 2004). *Distraction* is defined as attention, measured as eye movements, being captured by information that is irrelevant to the driving situation to the degree that a) insufficient attention is left for the primary control task of driving, *and/or* b) that driving performance (e.g. lane keeping or speed control) is compromised.

The real-time algorithm uses the same idea and procedure to localize the road center peak as was developed for the Percent Road Centre (PRC) measure in study II but defines the road-centre area as a circle with a 10 degrees radius instead of 8 as in studies II and III. During natural driving the road

centre peak gradually builds up, and after approximately one minute the road centre peak is stable and the classification starts.

Each gaze point is classified as being either “road-center” or “eyes-off-road” on the basis of whether it falls within the road centre area. Thus, the signal describes instantaneous or momentary distractions from what is happening in the forward direction. A consequence is that distractions can be towards in-vehicle sources such as information systems, but also towards external sources such as signs, scenery, or mirrors.

In addition to this momentary eyes-off-road (EOR) signal, three measures are calculated based on this signal a) PRC calculated as a running average using a one minute time window, b) single glance duration (SGD), and c) visual time sharing (VTS) calculated as a PRC running average using a 10 second time window. The algorithm can also be adaptive to different driving situations (see Engström & Victor, 2005; Larsson & Victor, 2005)

### **Visual distraction alert**

Many versions of a visual distraction alert were conceived (e.g. Victor, 2000a, 2000b, 2001a, 2001b, 2003) and tested in desktop, driving simulator, and on-road environments, including LEDs, icons, tones, seat vibration, and a recording of my two-year old son saying “look at the road”. The idea of the visual distraction alert is to help the a driver to realize that he/she is being ‘tricked’ into glancing away from the road too long or too often, and to ‘train’ him to recognize a limit. As such it was implemented as a preventative warning without direct coupling to driving performance deterioration.

In the first implementation (Figure 14 a)) five strings of LEDs were attached along the upper interior of the doors and the dashboard to create waves of light towards road centre. A later version used only three LEDs, see Figure 14 b). This three LED version was built to create the “visual rabbit illusion” discovered by Geldard (1975, 1982), whereby a running light stimulus is perceived as being localized between the 5 degree spaced locations if three flashes per position are used with a 100 msec interflash interval.

Different warning strategies were tested. A “tickle” warning, or a sparkling in the LEDs reflected in the windshield, was activated as a gentle pre-warning when PRC fell below 65% and when gaze was currently “eyes-off-road”. If the driver persisted in glancing away and PRC fell to 58% two waves of running lights along the five strings toward road centre were activated. The same waves of running lights were also activated when a single glance exceeded 3.4 seconds.

Following the feedback from 14 truck drivers who tested this warning on-road, the warning strategy was changed to be simpler by removing the tickle warning. 16 truck drivers then tested the simpler warning strategy on-road, giving it better ratings. The results of these experiments are unpublished as

yet. Similar visual distraction alerts were later developed and tested by Almén (2003) and Karlsson (2004).



Figure 14 a) and b). Examples of a visual distraction alert that redirects gaze using flashing LEDs, a) the first implementation used strings of smaller LEDs running along the doors and centre console towards the road centre. The LEDs at road centre were reflected on the windshield. Foto: Lars Ardarve. A later version used three strong LEDs as illustrated in b) with the last LED reflecting on the windshield. Foto: Volvo Truck Corp.

### **Cognitive distraction alert**

A cognitive distraction alert was conceived, developed in several versions, and tested in a driving simulator and on-road. The various versions of the alert flash LEDs reflected in the windshield in the centre and towards the visual periphery where signs appear. The latest version used three LEDs to flash at road centre, left, centre, right, and centre. The idea is to detect that a driver is cognitively distracted by measuring the gaze concentration effect associated with cognitive and auditory tasks found in study II and III. The cognitive distraction alert is issued when PRC reaches 92 %.

### **Hazard alerts – feedback on driving performance deterioration caused by distraction**

When a) visual or cognitive distraction is observed together with b) driving performance deterioration or c) potential hazards in the external environment (i.e. when eyes-off-road or visual time sharing is associated with safety hazards), more precise Hazard alerts can be activated (Victor, 2003). Hazard alerts are intended to redirect attention to potential hazards when drivers are distracted. For example, if steering-wheel movement or lane keeping behavior is deviant during secondary task glance behavior (VTS) then provide a hazard alert. Thus, hazard alerts take the visual distraction alert one step closer to potential collisions by detecting when driving control

task performance decrements or potential collisions occur simultaneously with distraction.

### Inattention sensitive driving support functions

Apart from direct inattention warnings, distraction detection (with the EOR, PRC, and VTS measures) can be used to make driving support functions inattention-sensitive (Victor, 2003, 2004). Examples of driving support functions are adaptive cruise control (ACC), forward collision warning (FCW), lane departure warning (LDWS). The idea is to enable/disable warnings depending on attention state, or set it functions into different modes. For example, PRC can be used to set a forward collision warning (FCW) system into 'sensitive' mode, and the instant eyes-on-road-centre signal could be used to decide weather a warning should be issued or not while a driver is operating a mobile telephone. They could also be used to adjust the time-gap (increase or decrease the safety distance) for an ACC system.

### Peripheral information displays

A system reducing the amount of in-vehicle glances is desirable from a safety standpoint because glances off-road are so highly tied to crashes (Neale et al., 2005). As described in Victor and Jarlengrip (2005a) displays can be adapted so that the driver can read or recognize information with peripheral vision. The idea is to present information on displays in such a way that eye movements towards that information and subsequent eye-fixations upon the information are not necessary. The information is readable or recognizable without having to move the eyes off the road. The information has to be presented simply and large enough to enable information extraction with peripheral vision.

A regression equation for visual acuity,  $y = 0.046x - 0.031$  deg, can be used to determine display characteristics (Anstis, 1974). Thus, for every degree of visual angle the minimum discriminable size increases by about 2.5 minutes of arc. Examples of information-types that could be presented are current speed, current road signs, navigation information, text messages, or gear information.

### Path keeping displays

As outlined in the introduction, driver has to perform two distinct tasks to control a vehicle's path: control the future path trajectory and control the present path position. Displays for supporting and improving the driver's performance in these two tasks are describes here and in Victor and Jarlengrip (2005b).

Given the significance of the far-path point for detecting future-path error, and the added priority drivers place on it when driving gets demanding, it follows that a system that assists the driver in detecting future-path error would be advantageous. If future-path error can be made more easily recognizable, then steering corrections improve, path control is improved, and the driver can place more priority on recognition and planning tasks because she is freer to move her eyes around to other objects and areas of vision during highly demanding situations.

### **Feedback to support future-path trajectory assessment**

To support the control of the future path trajectory, predictive information to the driver about the vehicle's actual future path can be presented enabling the driver to directly see the difference between where the vehicle is heading and how it compares to where the driver actually wants to go (i.e. making the error term more visible). This type of display ideally requires the system to have information of 1) head position and/or eye position, 2) a path prediction estimate, and 3) the means with which to present information. An example of a means to present information is a commercially available head up display (HUD), a diode laser, or a head mounted display.

### **Peripheral displays of lane-keeping information**

Control of the *present-path position* is achieved mainly by peripheral vision. In controlling the present-path position, the driver compares present position in path with desired present position in path and steers the vehicle to correct this error. The driver most often compares present position relative to lane markings, but can also regulate position relative to objects close to the vehicle.

Access to the information specifying present-path position is not always entirely accessible to our visual system. For example, when drivers operate information systems, such as a radio, peripheral information is blocked by the interior of the vehicle. If present-path error can be made more easily recognizable, then steering corrections become improved and unintentional lane exits can potentially be eliminated.

A number of peripheral displays of lane-keeping information are described in Victor and Jarlengrip (2005b). The idea is to display lane-keeping information in a simple enough way to be interpretable with peripheral vision only, when gaze is *not* directed at- or close to the lane-keeping display. The driver should not need to remove his or her eyes off the road. The display of lane-keeping information uses a combination of a lane tracker, a head/eye position sensor, and a means to present information.

In one embodiment, deviation in lane is represented by an increase in a line of light projected on the dashboard by a diode laser. The amount and

placement of the line of light being displayed corresponds to the amount of lane deviation (registered by the lane-tracking device) from a goal state. Goal-state markings are calculated from knowledge of head position and/or eye position, knowledge of the position in lane, and the width of the vehicle. The knowledge of head and eye position, and knowledge of the geometrics of the surfaces which the information is presented on, allows the system to position the goal-state markings to match a continuation of the lane or road markings.

In another embodiment, a horizontal array of LEDs (lightbar) is used to indicate lateral error from goal state. A similar concept has been developed to aid the control of farm machinery such as tractors, see Young and Mann (2002). Feedback from a peripheral vision lightbar is given based on position information from a GPS-based guidance system.

## Workload Management

Workload managers have been under development for some time now (see Engström, 2005a; Hoedemaeker, de Ridder, & Janssen, 2002). The general idea of a workload manager is to assist the driver in maintaining attention to driving by prioritizing system-initiated information and delaying presentation until driver workload is low. Workload refers to how busy a person is and the amount of effort they need to perform the tasks at hand. One further development suggested for workload managers is to help the driver manage telephone conversations (Engström, 2005b; Victor, 2003). Crundall, Bains, Chapman, and Underwood (2005) demonstrated that in-car conversations are suppressed by drivers and passengers during demanding traffic situations, whereas talking on the mobile telephone prevents suppression from taking place. Mobile phone conversations even encouraged drivers to make more utterances that they would normally do with a normal in-car conversation.

A potential solution to this problem is to have a workload manager pause spoken dialogue (e.g. telephone conversation) and system-initiated information (e.g. text-to-speech email, non-critical navigation system dialogue) during periods of high visual activity (Victor, 2003; Victor & Larsson, 2004). For example head and eye rotation variability can be used. One version of this was developed tested. It activated a pause tone and a visual message during periods of high visual activity while emails will be read out loud by a text-to-speech system.

This Attention Support section outlined a few practical ways, grounded in the theory and empirical studies outlined in this thesis, of potentially reducing crashes attributable to inattention. Perhaps they will inspire future research and development.

## Acknowledgements

Many thanks go to the Program Board for Swedish Automotive Research and Volvo Technology (especially Lars-Göran Rosengren) for shoeing this horse, the VISREC project for sending me off the general direction, and my talented trainer/supervisor Claes von Hofsten for leading me to theoretical waters and for teaching me how to perform some nice tricks.

Thankfully, my burdens were shared by Johan Engström, Petter Larsson, Emma Johansson, Dennis Saluäär, Alex Zelinsky and Seeing Machines, Peter Burns, Joanne Harbluk, Nina Åberg, Olle Blomberg, Anders Agnvall, Martin Magnusson, Christian Larsson, Christer Johansson, Johan Jarlengrip, Paul Piamonte, Karin Svensson, Werner Fröhling, Tracy Druce, and Tomas Heun. I thank you all, and the rest of the herd.

I am grateful for spending a fair bit of time wandering between different stables (Department of Mechanical Engineering, Department of Computer and Information Science, HMI Graduate School) and trainers (Erik Hollnagel, Håkan Alm, Kjell Ohlsson) at Linköping University before I went back home to the Department of Psychology at Uppsala University. I thank Claes von Hofsten, Kerstin Rosander, Erik Börjesson, and others for helping me feel so comfortable in the stable.

Three foals were born on the trip. Ossian, Rasmus, and Stella gave me ample experience with distraction, path-control disturbances, and competing high-level goals. I thank my mother and father Marie and Gunner Chistian-sen for good support and lineage. Most of all I thank my beloved mare Gunilla for a great trip.

I know I'm just being nostalgic; the age of the horse is over. It's time to drive a car like everyone else.

## References

- Allport, D. A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer and A. F. Sanders (Eds.), *Perspectives on Perception and action*. Hillsdale, NJ: Erlbaum.
- Allport, D. A. (1989). Visual attention. In M. I. Posner (Ed.), *Foundations of Cognitive Science*. Cambridge, MA: MIT Press.
- Allport, D. A. (1993). Attention and control: Have we been asking the wrong questions? In D. E. Meyer and S. Kornblum (Eds.), *Attention and Performance 14*. Cambridge, MA: MIT Press.
- Almén, L. (2003). *Reducing the effects of driver distraction: A comparison of distraction alerts on driver attention*. Licentiate thesis No. 1017, Linköping University, Linköping, Sweden.
- Ambinder, M. S. & Simons, D. J. (2005). Attention capture: the interplay of expectations, attention, and awareness. In L. Itti, G. Rees, and J. K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 69-75). Amsterdam: Elsevier.
- Anstis, S. M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research*, 14, 589-92.
- Arensberg, U. (2004). *Tiredness and distraction are major sources of danger in traffic – the truck monitors how you are feeling*. Press release 18 May, 2004. Volvo Truck Corporation.
- Atkinson, J. (2000). *The developing visual brain*. Oxford: Oxford University Press.
- Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, 40, 309-328.
- Baizer, J. S., Ungerleider, L. G., and Desimone, R. (1991). Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *Journal of Neuroscience*, 11, 168-190.
- Belt, B. L. (1969). Driver eye movements as a function of low alcohol concentrations. *Technical Report, Driving Research Laboratory*, Columbus, OH: Ohio State University.
- Billock, G., Koch, C., & Psaltis, D. (2005). Selective Attention as an Optimal Computational Strategy. In L. Itti, G. Rees, and J. K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 18-23). Amsterdam: Elsevier.
- Blakemore, S.-J., Frith, C. D., & Wolpert, D. W. (1999). Spatiotemporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience*, 11, 511-559.
- Brookhuis, K. A., de Vries, G., & de Ward, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention*, 23(4), 309–316.
- Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention. Bridging cognition and neurophysiology. *Psychological Review*, 112, 291-328.

- Carsten, O., Merat, J., Janssen, W., Johansson, E., Fowkes, M., & Brookhuis, K. A. (2005). *HASTE Final Report*. EU project HASTE, project no. GRD1/2000/25361.
- Castiello, U., Paulignan, Y., & Jeannerod, M. (1991). Temporal dissociation of motor responses and subjective awareness. *Brain*, 114, 2639-2655.
- Claesson, C. (2003). *Volvo trucks researches driver distraction*. Press release 5 May, 2003. Volvo Truck Corporation. Available: [http://www.volvo.com/trucks/global/en-gb/newsmedia/press\\_releases/](http://www.volvo.com/trucks/global/en-gb/newsmedia/press_releases/)
- Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., et al. (2000). A Common Network of Functional Areas for Attention and Eye Movements. *Neuron*, 21 (4), 761-773.
- Corbetta, M. & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*. 3 (3), 201-215.
- Craighero, L. & Rizzolatti, G. (2005). The premotor theory of attention. In L. Itti, G. Rees, and J. K. Tsotsos (Eds.), *Neurobiology of Attention*. (pp. 181-186). Amsterdam: Elsevier.
- Creem, S. H., & Proffitt, D. R. (2001). Grasping objects by their handles: A necessary interaction between cognition and action. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 218-228.
- Crundall, D., Bains, M., Chapman, P., & Underwood, G. (2005). Regulating conversation during driving: a problem for mobile telephones? *Transportation Research Part F*, 8 (3), 197-211.
- Crundall, D., Chapman, P., Phelps, N., & Underwood, G. (2003). Eye movements and hazard perception in police pursuit and emergency response driving. *Journal of experimental psychology: Applied*, 9 (3), 163-174.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193-222.
- Deyoe, E. A., & Brefczynski, J. (2005). Neurophysiological correlates of the attentional spotlight. In L. Itti, G. Rees, and J. K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 372-376). Amsterdam: Elsevier.
- Donges, E. (1978). A two-level model of driver steering behavior. *Human Factors*, 20(6), 691-707.
- Engström, J. (2005a). *Method and device for estimating workload for a driver of a vehicle*. European Patent Application EP 1 512 374 A1. European Patent Office.
- Engström, J. (2005b). *Method for handling and incoming call and vehicle on board arrangement*. European Patent Application EP 1 513 326 A1. European Patent Office.
- Engström, J. A., Johansson, E. K., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Psychology and Behaviour*, 8(2), 97-120.
- Engström, J. & Victor, T. (2005). *System and method for real-time recognition of driving patterns*. United States Patent No. US 6879969 B2.
- Ericsson, K. A., & Smith, J. (1991). *Toward a general theory of expertise: Prospects and limits*. Cambridge, UK: Cambridge university press.
- Evans, L. (2004). *Traffic Safety*. Bloomfield Hills, MI: Science Serving Society.
- Felleman, D. J., & van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, 1-47.
- Field, D. T., & Wann, J. P. (2005). Perceiving time to collision activates the sensorimotor cortex. *Current Biology*, 15(5), 453-8.
- Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Oxford: Oxford University Press.

- Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, 424, 769-771.
- Geldard, F. (1975). *Sensory saltation. Metastability in the perceptual world*. Hillsdale, NJ: Erlbaum.
- Geldard, F. A. (1982). Saltation in somesthesia. *Psychological Bulletin*, 92 (1), 136-175.
- Gibson J. J., (1958). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology*, 49, 182-194.
- Gibson, J. J., & Crooks, L. E. (1938). A theoretical Field-Analysis of Automobile-Driving. *American Journal of Psychology*, 51, 453-471.
- Goodale, M. A. (1998). Visuomotor control: Where does vision end and action begin? *Current Biology*, 8, 489-491.
- Goodale, M. A., & Humphrey, G. K. (1998). The objects of action and perception. *Cognition*, 67, 181-207.
- Goodale, M. A., & Murphy, K. (1997). Action and perception in the visual periphery. In P. Their, & H.-O. Karnath (Eds.), *Parietal Lobe Contributions to Orientation in 3D Space* (pp. 447-461). Heidelberg: Springer-Verlag.
- Goodale, M. A. & Milner, A. D. (2004). *Sight Unseen: An Exploration of Conscious and Unconscious Vision*. Oxford, UK: Oxford University Press.
- Goodale, M. A. & Westwood, D. A. (2004). An evolving view of duplex vision: Separate but interacting cortical pathways for perception and action. *Current Opinion in Neurobiology*, 14, 203-211.
- Goodale, M. A., Pélisson, D., & Prablanc, C. (1986). Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature*, 320, 748-750.
- Gredebäck, G. (2004). Infants' knowledge of occluded objects: Evidence of early spatiotemporal representations. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 134*, Uppsala: Acta Universitatis Upsaliensis.
- Green, M. (2000). "How long does it take to stop?" Methodological analysis of driver perception-brake times. *Transportation Human Factors*, 2 (3), 195-216.
- Green, P. (1999). *Visual and task demands of driver information systems* (Technical Report No. UMTRI-98-16). Ann Arbor, MI: The University of Michigan, Transportation Research Institute.
- Grigo, A., & Lappe, M. (1998). Interaction of stereo vision and optic flow processing revealed by an illusory stimulus. *Vision Research*, 38, 281-290.
- Groos, K. (1896). *Die Spiele der Thiere*. Fischer, Jena.
- Gu, J., Meng, M., Cook, A., & Faulkner, M. G. (2000). Analysis of Eye Tracking Movements Using FIR Median Hybrid Filters. In *Eye Tracking Research & Applications Symposium 2000*. Palm Beach Gardens, FL, USA.
- Hammel, K. R., Fisher, D. L., & Pradhan, A. K. (2002). Verbal and spatial loading effects on eye movements in driving simulators: A comparison to real world driving. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 2174-2178).
- Harbluk, J.L., Noy, Y.I., & Eizenman, M. (2002). The Impact of Cognitive Distraction on Driver Visual Behaviour and Vehicle Control. *Transport Canada Technical Report*. Available at: <http://www.tc.gc.ca/roadsafety/tp/tp13889/en/menu.htm>
- Hoedemaeker, M., de Ridder, S. N., & Janssen, W. H. (2002). *Review of European Human Factors Research on Adaptive Interface Technologies for Automobiles*. TNO-report. TM-02-C031.

- Hollands, M. A., Marple-Hovatt, D. E., Henkes, S., & Rowan, A. K. (1995). Human eye movements during visually guided stepping. *Journal of Motor Behavior*, 27, 155-163.
- Horrey, W. J. & Wickens, C. D. (2004). The impact of cell phone conversations on driving: A meta-analytic approach. (Technical Report AHFD-04-2/GM-04-1). Warren, MI: General Motors Corporation.
- Ingle, D. J. (1973). Two visual systems in the frog. *Science*, 181, 1053-1055
- ISO 15007-1 *Road vehicles – Measurement of driver visual behaviour with respect to transport information and control systems Part 1: Definitions and parameters*. TC22/SC 13, International Organization for Standardization.
- ISO 15007-2 *Road vehicles – Measurement of driver visual behaviour with respect to transport information and control systems Part 2: Equipment and procedures*. TC22/SC 13, International Organization for Standardization.
- Itti, L. (2005). Models of Bottom-up Attention and Saliency. In L. Itti, G. Rees, and J. K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 576-582). Amsterdam: Elsevier.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40, 1489-1506.
- Itti, L., Rees, G., & Tsotsos, J. K. (2005). *Neurobiology of Attention*. Amsterdam: Elsevier.
- Ito, T. & Miki, Y. (1997). Japan's Safety Guideline on In-Vehicle Display Systems. *Proceedings of the Fourth ITS World Congress*, Brussels, Belgium: VERTIS, CD-ROM.
- James, W. (1890). *Principles of Psychology, Volume 2*. New York: Henry-Holt.
- Janelle, C. M., Singer, R. N., & Williams, A. M. (1999). External distraction and attention narrowing: visual search evidence. *Journal of Sport and Exercise Psychology*, 21, 70-91.
- Johansson, E., Carsten, O., Janssen, W., Jamson, S., Jamson, H., Merat, J. et al. (2005). *HMI and Safety-Related Driver Performance, Deliverable 3*. EU project HASTE, project no. GRD1/2000/25361.
- Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye-Hand Coordination in Object Manipulation. *The Journal of Neuroscience*, 21(17), 6917-6932.
- Jonsson, B., & von Hofsten, C. (2003). Infants ability to track and reach for temporarily occluded objects. *Developmental Science*, 6, 88-101.
- Kalat, J. W. (2004). *Biological Psychology*. Belmont, CA: Thomson Wadsworth.
- Kaluger, N. A. & Smith, G. L., Jr. (1970). Driver eye movement patterns under conditions of prolonged driving and sleep deprivation. Night Visibility and Driver Behaviour. *Highway Research Record*, No. 336, Washington, D.C.
- Karlsson, R. (2004). *Evaluating driver distraction countermeasures*. Masters thesis. Linköping University, Linköping, Sweden.
- Koch, C. (2004). *The quest for consciousness. A neurobiological approach*. Englewood: Roberts and Company Publishers.
- Koch, C. & Crick, F. (2001). On the zombie within. *Nature*, 411, 893.
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4, 219-227.
- Kowler, E. (1990). The role of visual and cognitive processes in the control of eye movement. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 1-47). Amsterdam: Elsevier.
- Lamble, D., Laakso, M., & Summala, H. (1999). Detection thresholds in car following situations and peripheral vision: Implications for positioning of visually demanding in-car displays. *Ergonomics*, 42 (6), 807-815.

- Lamme, V. A. F. (2005). The difference between visual attention and awareness: A cognitive science perspective. In L. Itti, G. Rees, and J. K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 167-174). Amsterdam: Elsevier.
- Land, M. F. (1992) Predictable eye-head coordination during driving, *Nature*, 359.
- Land, M. (1998). The visual control of steering. In L. R. Harris and M. Jenkin (Eds.), *Vision and Action* (pp. 163-180). Cambridge, UK: Cambridge University Press.
- Land, M., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society (London)*, 352, 1231-1239.
- Land, M. F. & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, 41, 3559-3565.
- Land, M., & Horwood, J. (1995). Which parts of the road guide steering? *Nature*, 377, 339-340.
- Land, M.F., & Lee, D.N. (1994). Where we look when we steer. *Nature*, 369, 742-744.
- Lansdown, T. (2001). Causes, Measures, and Effects of Driver Visual Workload. In P. Hancock & P. Desmond (Eds), *Stress, Workload, and Fatigue*. New Jersey: Erlbaum.
- Lappe, M., Bremmer, F., & van den Berg, A. V. (1999). Perception of self-motion from visual flow. *Trends in Cognitive Sciences*, 3(9), 329-336.
- Larsson, P. (2003). *Automatic Visual Behavior Analysis*. (ISRN LITH-ISY-EX-3259-2002). Master's thesis, Linköping University, Linköping, Sweden.
- Larsson, P. & Victor, T. W. (2005). *Method and arrangement for interpreting a subjects head and eye activity*. United States Patent Application Publication. Pub. No.: US2005/0073136 A1.
- Lee, D. N., & Lishman, R. (1977). Visual control of locomotion. *Scandinavian Journal of Psychology*, 18, 224-230.
- Leigh, R. J., & Zee, D. S (1999). *The neurology of eye movements*. Oxford: Oxford University Press.
- Logan, G. D. (2002). An instance theory of attention and memory. *Psychological Review*, 109, 376-400.
- Maciejowski, J. M. (2002). *Predictive Control with Constraints*. Harlow, UK: Prentice-Hall, Pearson Education.
- Mack, A., & Rock, I. (1998). *Inattentional Blindness*. Cambridge, MA: MIT Press.
- Marr D. (1982). *Vision*. San Fransisco, CA: W. H. Freeman.
- McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience*, 19, 431-441.
- Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.
- Milner, A. D., and Goodale, M. A. (1998). The visual brain in action. Précis of: A. David Milner and Melvyn A. Goodale *The Visual Brain in Action* (Oxford Psychology Series, No. 27). Oxford University Press. Available: <http://psyche.cs.monash.edu.au/v4/psyche-4-12-milner.html>
- Most, S. B., Scholl, B. J., Clifford, E., & Simons, D. J. (2005) What you see is what you set: Sustained inattentional blindness and the capture of awareness. *Psychological Review*, 112(1), 217-242.
- Muira, T. (1990). Active function of eye movement and useful field of view in a realistic setting. In R. Groner, G. dYdewalle, R. Parham (Eds.) *From Eye to Mind: Information Acquisition in Perception, Search and Reading*. Elsevier Science Publishers B.V. (North-Holland).

- Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences*, 5(7), 309-315.
- Näätänen, R. & Summala, H. (1976). *Road user behaviour and traffic accidents*. New York: North Holland.
- Neale, V. L., Dingus, T. A., Klauer, S. G., Sudweeks, J., and Goodman, M. (2005). An overview of the 100-car naturalistic study and findings. *Paper presented at the 19th International Technical Conference on Enhanced Safety of Vehicles (ESV)*, Washington, DC, June 6-9, 2005.
- Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer and A. F. Sanders (Eds.), *Perspectives on Perception and Action*. Hillsdale, NJ: Erlbaum.
- Neumann, O. (1990). Visual attention and action. In O. Neumann and W. Prinz (Eds.), *Relationships between Perception and Action: Current Approaches*. Berlin: Springer.
- Neumann, O., & Sanders, A. F. (1996). *Handbook of Perception and Action. Volume three: Attention*. San Diego: Academic Press.
- Nilsson, R. (2001). Safety Margins in the Driver. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 106*, Uppsala: Acta Universitatis Upsaliensis.
- Nunes, L., & Recarte, M. A. (2002). Cognitive demands of hands-free-phone conversation while driving. *Transportation Research Part F*, 5, 133-144.
- Östlund, J., Nilsson, L., Carsten, O., Merat, H., Jamson, H., Jamson, S. et al. (2004). *HMI and Safety-Related Driver Performance, Deliverable 2*. EU project HASTE, project no. GRD1/2000/25361.
- Patla, A. E., & Vickers, J. N., (1997). Where and when do we look as we approach and step over an obstacle in the travel path? *NeuroReport*, 8, 3661-3665.
- Pavel, M. (1990) Predictive control of eye movement. In Kowler, E. (Ed.) *Eye movements and their role in visual and cognitive processes* (pp. 71-113). Amsterdam: Elsevier.
- Posner, M. I., (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.
- Proffitt, D. R., Bhalla, M., Gossweiler, R., & Midgett, J. (1995). Perceiving geographical slant. *Psychonomic Bulletin & Review*, 2, 409-428.
- Recarte, M. A., & Nunes, L. M. (2000). Effects of Verbal and Spatial-Imagery Tasks on Eye Fixations While Driving. *Journal of Experimental Psychology: Applied*, 6 (1), 31-43.
- Recarte, M. A., & Nunes, L. M. (2003). Mental Workload While Driving: Effects on Visual Search, Discrimination and Decision Making. *Journal of Experimental Psychology: Applied*, 9 (2), 119-137.
- O'Regan, J. K. & Noë, A., (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5), 939-1011.
- Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition*, 7 (1/2/3), 17-42.
- Rensink, R. A. (2002a). Change Detection. *Annual Review of Psychology*, 53, 245-77.
- Rensink, R. A. (2002b). Visual Attention. In L Nadel (Ed.), *Encyclopedia of Cognitive Science*. London: Nature Publishing Group.
- Reynolds, J. H., & Desimone, R. (1999). The role of neural mechanisms of attention in solving the binding problem. *Neuron*, 24, 19-29.
- Rizzolatti, G., Riggio, L., Dascola, I., & Umiltà, C. (1987). Reorienting attention across the horizontal and vertical meridians – evidence in favor of a premotor theory of attention. *Neuropsychologia*, 25, 31-40.

- Rockwell, T. H. (1972). Eye movement analysis of visual information acquisition in driving: An overview. *Proceedings of the 6th Biennial Conference of the Australian Road Research Board* (pp. 316-331). August 1972, Canberra.
- Rossetti, Y., & Pisella, L. (2002). Several 'vision for action' systems: A guide to dissociating and integrating dorsal and ventral functions. In W. Prinz & B. Hommel (Eds.), *Attention and Performance XIX: Common mechanisms in perception and action* (pp. 62-119). Oxford, UK: Oxford University Press.
- Rumar, K. (1990). The basic driver error: late detection. *Ergonomics*, *33* (10/11), 1281-1290.
- Salvucci, D. D., Boer, E. R., & Liu, A. (2001). Toward an integrated model of driver behavior in a cognitive architecture. *Transportation Research Record*, *1779*, 9-16.
- Salvucci, D. & Goldberg, J. (2000). Identifying Fixations and Saccades in Eye-tracking Protocols. *Proceedings of the Eye Tracking Research & Applications Symposium 2000*. Palm Beach Gardens, FL, USA.
- Salvucci, D. D., & Gray, R. (2004). A two-point visual control model of steering. *Perception*, *33*(10), 1233-48.
- Senders, J. W., Kristofferson, A.B., Levison, W. H., Dietrich, C. W., & Ward, J. L. (1967). The attentional demand of automobile driving. *Highway Research Record*, *195*, 15-33.
- Seppelt, B., & Wickens, C.D. (2003). *In-vehicle tasks: Effects of modality, driving relevance, and redundancy*. (Technical Report AHFD-03-16/GM-03-2). Aviation Human Factors Division, University of Illinois at Urbana-Champaign.
- Shelton, L.R. (2001). *Statement on Driver Distractions before the Subcommittee on Highway and Transit Committee on Transportation and Infrastructure, U.S. House of Representatives*, May 9, 2001. Available: <http://www.nhtsa.dot.gov/nhtsa/announce/testimony/distractiontestimony.html>.
- Shmuelof, L., & Zohary, E. (2005). Dissociation between Ventral and Dorsal fMRI Activation during Object and Action Recognition. *Neuron*, *47*, 457-470.
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattention blindness for dynamic events. *Perception*, *28*, 1059-1074.
- Somers, D. C., & McMains, S. A. (2005). Spatially-specific attentional modulation revealed by fMRI. In L. Itti, G. Rees, and J. K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 377-382). Amsterdam: Elsevier.
- Stelmach, L. B., Bourassa, C. M., & di Lollo, V. (1984). Detection of stimulus change: The hypothetical roles of visual transient responses. *Perception and Psychophysics*, *35* (3), 245-255.
- Summala, H., Nieminen, T. & Punto, M. (1996). Maintaining lane position with peripheral vision during in-vehicle tasks. *Human Factors*, *38* (3), 442-451.
- Tijerina, L., Kiger, S., Rockwell, T., & Tornow, C. (1995). *NHTSA heavy vehicle driver workload assessment final report supplement -- Workload assessment of in-cab text messaging system and cellular phone use by heavy vehicle drivers on the road*. (Technical Report DOT HS 808 467 (7A)). Washington, DC: National Highway Traffic Safety Administration.
- Treat, J. R., Tumbas, N. S., McDonald, S. T., Shinar, D., Hume, R. D., Mayer, R. E., Stansifer, R. L., & Castellan, N. J. (1977). *Tri-Level Study of the Causes of Traffic Accidents: Final Report – Executive Summary*. (Technical Report DOT HS 805 099). US Washington D.C.: Department of Transportation, National Highway Traffic Safety Administration.
- Tsotsos, J. K., Itti, L., & Rees, G. (2005). A brief and selective history of attention. In L. Itti, G. Rees, and J. K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. xxii-xxxii). Amsterdam: Elsevier.

- Underleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, Goodale, M. A., & Mansfield, R. J. W. (Eds.), *Analysis of Visual Behavior*, (pp. 549-586). Cambridge, MA: MIT Press.
- Victor, T. (2000a). On the need for attention support systems. *Journal of Traffic Medicine*. Vol. 28, No. 2S.
- Victor, T. (2000b). A Technical Platform for Inattention Research. *Proc Internet Forum on The Safety Impact of Driver Distraction When Using In-Vehicle Technologies*, (NHTSA, DOT, Washington DC), Available: [www-nrd.nhtsa.dot.gov/departments/nrd-13/driver-distraction/Welcome.htm](http://www-nrd.nhtsa.dot.gov/departments/nrd-13/driver-distraction/Welcome.htm).
- Victor, T. (2001a). Envisioning attention support systems. Presented at the Graduate School for Human Machine Interaction workshop, Linköping, Sweden, February 2, 2001.
- Victor, T. (2001b). *Attention Management HMI – Functional Requirements for Iteration 1*. (Engineering Report ER-520081). Volvo Technology Corporation, Göteborg, Sweden.
- Victor, T. (2003). *System and method for monitoring and managing driver attention loads*. United States Patent Application Publication. Pub. No.: US2003/0181822 A1.
- Victor, T. (2004). *Method and arrangement for controlling vehicular subsystems based on interpreted driver activity*. International application published under the patent cooperation treaty. International publication number PCT WO 2004/108466 A1. International Bureau: World Intellectual Property Organization.
- Victor, T. W., & Åberg, N. (2005). *The driving performance cost of gaze eccentricity in full attention and cognitive load conditions*. Manuscript. Department of Psychology, Uppsala University, Sweden.
- Victor, T. W., & Jarlengrip, J. (2005a). *Method and system for presenting information*. International application published under the patent cooperation treaty. International publication number PCT WO 2005/054786 A1. International Bureau: World Intellectual Property Organization.
- Victor, T. W., & Jarlengrip, J. (2005b). *Method and system for supporting path control*. International application published under the patent cooperation treaty. International publication number PCT WO 2005/053991 A1. International Bureau: World Intellectual Property Organization.
- Victor, T. W. & Larsson, P. (2004). *Method and arrangement for interpreting a subjects head and eye activity*. International application published under the patent cooperation treaty. International publication number PCT WO 2004/034905 A1. International Bureau: World Intellectual Property Organization.
- Victor, T. W. & Larsson, P. (2005). *Method and system for recognizing driver impairment*. International application published under the patent cooperation treaty. International publication number PCT WO 2005/052878 A1. International Bureau: World Intellectual Property Organization.
- van der Heijden, A. H. C. (1996). Visual Attention. In O. Neumann, & A. F. Sanders (Eds.), *Handbook of Perception and Action. Volume three: Attention*. (pp. 5-42). San Diego: Academic Press.
- von Hofsten, C. (1993). Prospective control: A basic aspect of action development. *Human Development*, 36, 253-270.
- von Hofsten, C. (1995). Planning and perceiving what is going to happen next. In M. Haith, J. Benson, R. Roberts, & B. Pennington (Eds.), *The development of future oriented processes* (pp. 63-86). Chicago: University of Chicago press.
- von Hofsten, C. (2004). An action perspective on motor development. *Trends in Cognitive Sciences*, 8 (6), 266-272.

- Wann, J. P., & Land, M. F. (2000). Steering with or without the flow: is the retrieval of heading necessary? *Trends in Cognitive Sciences*, 4 (8), 319-324.
- Wann, J. P., & Swapp, D. K., (2000). Why you should look where you are going. *Nature Neuroscience*, 3 (7), 647 – 648.
- Wann J. P., Swapp D. K., & Rushton S. K. (2000). Heading Perception and the allocation of attention. *Vision Research*, 40, 2533-2543.
- Wann, J. P., & Wilkie, R. M. (2004). How do we control high speed steering? In L. M. Vaina, S. A. Beardsley, S. K. Rushton (Eds.), *Optic Flow and Beyond* (pp.371-389). Dordrecht: Kluwer Academic Publishers.
- Weiskrantz, L., Elliot, J., Darlington, C. (1971). Preliminary observations of tickling oneself. *Nature*, 230, 598-599.
- Wierwille, W. W. (1993a). An initial model of visual sampling of in-car displays and controls. In A. G. Gale, I. D. Brown, C. M. Haslegrave, H. W. Kruysse, and S. P. Taylor (Eds.), *Vision in Vehicles IV* (pp. 271-282). Amsterdam: North-Holland.
- Wierwille, W. W. (1993b). Visual and manual demands of in-car controls and displays. In B. Peacock and W. Karwowski (eds.), *Automotive Ergonomics* (pp. 299-320). London: Taylor and Francis.
- Wierwille, W. W., Antin, J. F., Dingus, T. A., & Hulse, M. C. (1988). Visual attentional demand of an in-car navigation display system. In A. G. Gale, M. H. Freeman, C. M. Haslegrave, P. Smith, and S. P. Taylor (Eds.), *Vision in Vehicles II* (pp. 307-316). Amsterdam: North-Holland.
- Wierwille, W. W. & Tijerina, L. (1998). Modelling the relationship between driver in-vehicle visual demands and accident occurrence. In A. G. Gale, I. D. Brown, C. M. Haslegrave, and S. P. Taylor (Eds.), *Vision in Vehicles VI* (pp. 233-244). Amsterdam: North-Holland.
- Wilkie, R. M. & Wann, J. P. (2002). Driving as Night Falls: The Contribution of Retinal Flow and Visual Direction to the Control of Steering. *Current Biology*, 12, 2014-2017.
- Wilkie, R. M., & Wann, J. P. (2003a). Controlling Steering and Judging Heading: Retinal flow, visual direction and extra-retinal information. *Journal of Experimental Psychology: Human Perception and Performance*, 29 (2), 363 - 378.
- Wilkie, R. M., & Wann, J. P. (2003b) Eye-movements aid the control of locomotion. *Journal of Vision*, 3, 1-3.
- Wilkie, R. M. & Wann, J. P. (2005). The Role of Visual and Non-Visual Information in the Control of Locomotion. *Journal of Experimental Psychology: Human Perception and Performance*. 31(5).
- Wolpert, D. M., Miall, R. C., Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*. 2 (9), 338-347.
- Young, S., & Mann, D. D. (2002). Development and evaluation of peripheral display for agricultural guidance aids. Paper presented at AIC2002: The Canadian society for engineering in agricultural, food and biological systems. Available: <http://www.engr.usask.ca/societies/csae/PapersAIC2002/CSAE02-204.pdf>
- Zhang, H. & Smith, M. (2004a). A final report of SAfety VEhicles using adaptive Interface Technology (Phase I: Task 7): A Literature Review of Visual Distraction Research. Available: <http://www.volpe.dot.gov/opsad/saveit/docs.html>
- Zhang, H. & Smith, M. (2004b). A final report of SAfety VEhicles using adaptive Interface Technology (Phase I: Task 7): Visual Distraction Research. Available: <http://www.volpe.dot.gov/opsad/saveit/docs.html>
- Zwahlen, H.T., Adams, C.C. Jr., & de Bald, D.P. (1988). Safety Aspects of CRT Touch Panel Controls in Automobiles. In A.G. Gale et al. (Eds.) *Vision in Vehicles II*, (pp. 335-344). The Netherlands: North Holland.

# Acta Universitatis Upsaliensis

*Digital Comprehensive Summaries of Uppsala Dissertations  
from the Faculty of Social Sciences 9*

Editor: The Dean of the Faculty of Social Sciences

A doctoral dissertation from the Faculty of Social Sciences, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences”.)

Distribution: [publications.uu.se](http://publications.uu.se)  
urn:nbn:se:uu:diva-6241



ACTA  
UNIVERSITATIS  
UPSALIENSIS  
UPPSALA  
2005