ATTENTION SELECTION AND TASK INTERFERENCE IN DRIVING: AN ACTION-ORIENTED VIEW

Johan Engström*, Gustav Markkula and Trent Victor
Volvo Technology Corporation/SAFER, Chalmers University of Technology

Götaverksgatan 10
405 08 Göteborg, Sweden
Tel. +46 31 322 27 80
johan.a.engstrom@volvo.com

ABSTRACT

A major limitation in existing approaches to attention selection in driving research is a lack of focus on driving as an active process which is strongly driven by task goals and context. This paper outlines an attention selection model which is action-oriented and intended to capture real world driving situations. Based on this model, three main types of task interference in driving are proposed. Implications of the action-oriented view for the design of driver behaviour studies and automotive human machine interfaces are discussed.

INTRODUCTION

Driving is a complex task that requires a great deal of behavioural flexibility to manage and adapt to rapidly changing situational demands. A key aspect of this flexibility is to select the stimuli and actions that are relevant to the driving context and task goals, while ignoring others. This selection process is generally referred to as *attention*, although there is still no scientific consensus on a detailed definition of the term.

Drivers' attention selection is governed by a multitude of factors. For example, consider the following scenario, further illustrated in Figure 1. A driver, heavily engaged in a cell phone conversation, is driving on a rural road. After a while he approaches a busy T-junction intersected by a two-way bicycle lane. The T-junction is signalised by a road sign a few hundred meters earlier. In addition, a tree occludes the bicycle lane to the right. When the driver approaches the junction, a bicyclist approaches from the right on collision course with the driver.

Which factors determine whether the bicyclist will capture the driver's attention in order to elicit timely braking? First, *stimulus saliency*, determined e.g. by the colour of the bicyclist's clothes, reflexes, background contrast, movement etc., is an important factor determining attention capture [14]. The visual input may also be masked by temporary *occlusions*, e.g. blinks, saccades or, in the present example, the tree. Moreover, detection thresholds increase with increased *visual eccentricity*, that is, the angle of the stimulus relative to the gaze direction [18], up to the point where detection performance becomes zero at the boundary of the field of view. Hence, detection will be less effective if the bicyclist first appears in the peripheral, compared to the central, field of view (i.e. if the driver is looking away). Moreover, research on *change blindness* [22, 36] has demonstrated that people are notoriously bad at detecting changes when the transient change (e.g. the stimulus onset) does not impinge on the retina. Hence, if, in the present example, the bicyclist's appearance behind the tree is masked by a blink or occurs outside the visual field of view, attention capture would be expected to be severely impaired.

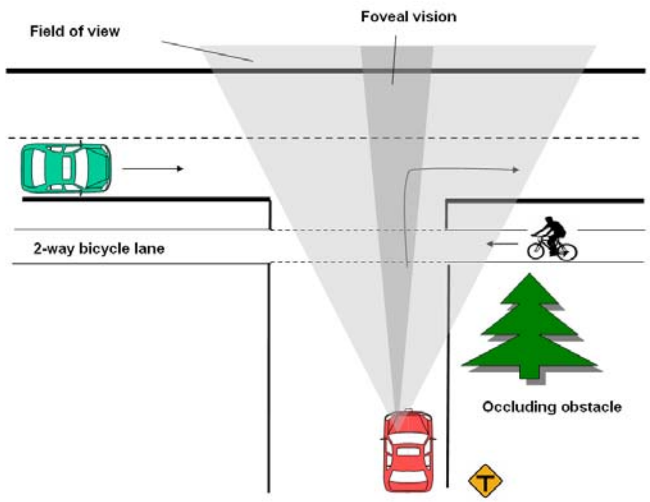

Figure 1 Example scenario: A driver is approaching a T-junction with an intersecting two-way bicycle lane, where a bicyclist is approaching from the right.

However, drivers' attention selection is also strongly driven by *expectations* on how the upcoming situation will unfold. As shown by Summala and colleagues [34, 42], in situations such as turning right at T-junctions and roundabouts, where the main expected hazards are cars coming from the left, drivers tend to focus their visual scanning to the left visual field. As a consequence, the bicyclist in our example, appearing in the right field of view, may fail to capture attention simply because he appears outside the field of view. Furthermore, expectancy may drive the selection of *attention*, irrespectively of gaze allocation, towards certain locations and/or objects. Laboratory studies have shown that when attention is allocated to a strongly *perceptually loading* task (for example, a detection task where the contrast between the target stimulus and surrounding distractor stimuli is small), the ability of peripheral stimuli irrelevant to that task to capture attention is strongly reduced [19, 20]. In fact, people may even fail to detect salient stimuli appearing right before the eyes when engaged in perceptually demanding tasks, as demonstrated by studies in the *inattentional blindness* paradigm [24, 40]. Thus, in our example, the fact that our driver's visual attention is strongly driven by expectations on *cars* to *the left* may further reduce the chance that the bicyclist will capture the driver's attention. Finally, non-visual tasks that put high demands on working memory, such as phone conversation, may also impair object/event detection/recognition [41]. Hence, in the present scenario, the phone conversation may further impair the ability to detect the bicyclist.

As illustrated by this example, attention selection is driven both *bottom-up*, as determined e.g. by stimulus saliency, temporary occlusions and visual eccentricity, and *top-down*, as driven by expectations relating to the current driving context and goals. Importantly, top-down attention selection does not necessarily have to involve effort or conscious awareness. Rather, routine situations, (e.g. overtaking on a sparsely trafficked motorway) are often negotiated relatively effortless, but may still involve elements of top-down, proactive, and context-dependent selection, based on “implicit” expectations on how the situation will develop. However, more effortful processes associated with conscious awareness may come into play in non-routine or inherently difficult situations, such as negotiating a complex intersection. Models of drivers’ attention selection must be able to account for both bottom-up and top-down factors and the dynamic interaction between them in both routine and non-routine situations.

Contemporary “standard” information processing (IP) models of operator performance are based on the common assumption of a unidirectional sequence of distinct information processing stages, e.g. perception, decision making and response (see e.g. [45] for a general review of IP models). In this paradigm, attention selection is generally viewed as a consequence of *limited capacity* in the information processing system. However, there is still substantial debate over the nature of this limited capacity. One aspect of this debate concerns whether selection occurs early [3, 12] or late [27] in the processing chain. Another concerns whether the limitations are best conceptualised as due to processing “bottlenecks” [32, 43] or competition for shared resources [17, 44]. Yet another issue of dispute is whether capacity is limited by a single resource/bottleneck [17, 32, 43] or whether there are multiple resources, each with its own capacity limitations [44].

In the IP tradition, the key experimental method for investigating possible bottlenecks and/or capacity limitations has been the *dual task paradigm*, where subjects are instructed to perform two tasks simultaneously. While this approach is potentially useful for studying the *momentary* interference between different types of tasks (i.e. the costs of concurrence during the moments in time when the two tasks are attempted simultaneously; see e.g. [23] for an application in the driving domain), most IP models have little to say about the mechanisms behind proactive, goal- and context-dependent, attention allocation exemplified above. For example, while IP models addresses why the response to the bicyclist in the T-junction scenario above is slowed by the phone conversation[1], they do not account for why the driver chose to engage in the conversation in the first place, or how he could manage the conversation (e.g. by interruption) based on expectations on how the situation will unfold.

Another limitation of traditional IP models is that they mainly apply to “interference in the brain”, including what Wickens [44] refers to as *structural interference*, related e.g. to “the difficulties in simultaneously performing two independent motor acts (e.g. rubbing the head and patting the stomach)” (p. 71) and more general capacity limitations related to resource sharing [16]. However, issues related to what Wickens [44] refers to as *peripheral interference*, i.e. physical constraints such as the fact that “the eye cannot view two separated locations at once, nor can a given finger simultaneously depress two keys, or the mouth utter two words at once” (p. 71), has been largely ignored or sometimes confounded with structural interference. Wickens [44] explicitly states that “structural interference is compatible with resource theory; peripheral interference is not” (p. 71). While effects of peripheral interference may be trivial in laboratory dual task studies, they are of key significance for attention selection in real-world driving situations. For example, traditional IP models does not seem to account for impaired detection of the bicyclist in our example due to the bicyclist

[1] However, there stills seems to be little consensus on how effects of phone use on driving performance should be interpreted in terms of IP models (see e.g. the debate between Moray (1999) and Wickens (1999)).

appearing outside the field of view, and the potential further consequences in terms of change blindness discussed above.[2]

Finally, as pointed out by Neumann [30], while traditional IP models have mainly focused on locating the attentional bottleneck in the chain of processing stages or identifying the set of limited resources that best fit empirical dual task data, they do not provide any detailed accounts of the actual mechanisms behind attention selection.

In sum, traditional attention models appear insufficient to provide a unified account of the different bottom-up and top-down factors that govern attention selection and task interference in real-world driving situations such as the T-junction scenario exemplified above. The objective of the present paper is to outline some initial ideas towards an alternative, action-oriented, model of attention selection and task interference in driving which better accounts for proactive, goal- and context-driven attention allocation in the real world.

A CONCEPTUAL MODEL

The key starting point for the proposed model is that attention selection in driving can be viewed in terms of action selection and, in the context of driving, as part of a larger repertoire of adaptive behaviours (e.g. slowing down, adjust lateral position) with the general purpose maintain sufficient perceived *safety margins* to potential obstacles and other hazards [10, 11]. Action-oriented views of attention selection have been developed in parallel to the information processing models reviewed above [1, 29, 30, 31] but have so far not been widely applied in the driving domain. While, as reviewed above, traditional IP models tend to view attention selection as a consequence of limited capacity, the present action oriented perspective views *selection*, in particular the selection of action, as the primary phenomenon of interest.

The proposed model contains three main components: (1) *Sensory and effector systems* interacting with the environment, (2) competing and cooperating *schemata*, implementing routine actions and action patterns and (3) *supervisory control*, which may be used to bias the schemata when demanded by the task at hand. The model is strongly related to the model by Norman and Shallice [31] and the more detailed computational implementations by Cooper, Shallice and colleagues [6, 7], although it differs from these models in the details. Earlier versions of the present model have been presented in [9, 10]. The general organisation of the model is illustrated in Figure 2. The three main model components are further described in the following sections.

[2] To address this limitation, Wickens and Horrey [47] have recently proposed a visual scanning model known as SEEV (Saliency, Effort, Expectancy and Value), intended as a complement to Wickens' original Multiple Resource Theory (MRT) [43]. More specifically, SEEV is invoked to explain phenomena related to change blindness while MRT should account for inattentional blindness [47]. While the SEEV model may be valuable for predicting gaze distributions in applied contexts, it does not provide any detailed account for the actual mechanisms that generate these distributions. In particular, it does not account for how bottom-up and top-down factors interact.

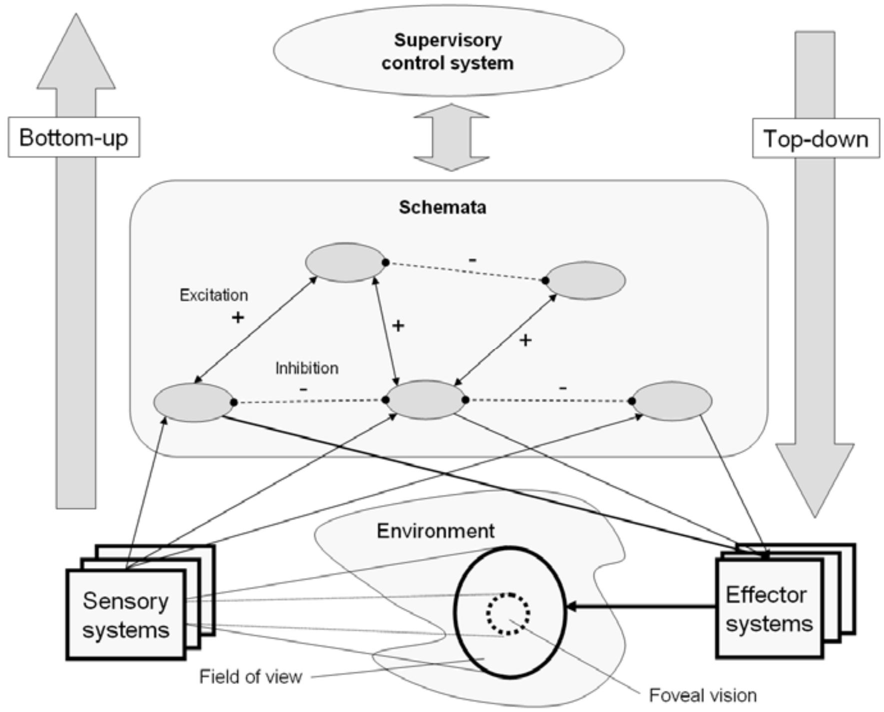

Figure 2 General organisational principles of the model. Solid lines represent excitatory links and dashed line inhibitory links.

Sensory and effector systems

In line with an embodied view on cognition, the physical interaction with the environment is assumed to be of fundamental importance and the basis for all processes at higher levels in the model. In the present model, this interaction is effected by closely coupled sensory and effector systems, as further illustrated at the bottom of Figure 2.

The sensory systems refer to the sensor organs, e.g. the retina, the cochlea, the skin sensors etc. which provide input to the schemata. Here, we will focus on the visual sensory modality as this is naturally the most relevant for driving. Physical stimulation of retinal photoreceptors yields visual sensory input. This input is fundamentally limited by the physical range of the sensors, in particular the frequency range and the spatial field of view. Visual performance generally degrades with retinal eccentricity and high acuity vision is only available at a small region in the centre of the retina known as the fovea. The effective strength of the (visual) sensory input is assumed to depend on (1) stimulus properties (e.g. intensity background contrast, abrupt onset and movement), (2) temporary visual occlusions (e.g. blinks or saccades) and (3) the visual angle of the stimulus relative to the retina.

The effector systems refer to muscles controlling the limbs and the sensory systems. Their role in action is both to change the state-of-affairs in the environment and to actively control the sensory systems (e.g. by means of eye, head or whole body movements). For example, the role of an eye movement action is generally to bring stimuli within the field of view, thus actively creating visual input for further actions.

Schemata

For present purposes, a *schema* is defined as a functional unit that represents a simple action or a more complex action pattern. Schemata may also represent processes for abstract though operation not directly linked to the sensory-motor surfaces (in line with e.g. [37]). The functional nature of schemata implies that they may be related in a many-to-many way to the neural substrate that realises them physically. Thus, a schema may involve several different brain areas, and a particular brain substrate (e.g. a population of neurons) may be involved in realising several different schemata. Schema theory provides a useful abstract language for the study of action-oriented perception, which can be naturally linked to brain theory as well as cognitive modelling and robotics [2]. The schema level in the present model roughly corresponds to the contention scheduling system proposed by Norman and Shallice [31]. The proposed selection mechanism is also essentially similar to the widely accepted *biased competition* model of attention selection at the neural level [8, 26] although conceptualised here on a more abstract level and from an action-oriented perspective.

In the present model, schemata are hierarchically organised. Schemata at the lowest level are directly linked to the sensory and motor surfaces and implement basic, real-time, sensory-motor acts such as steering to keep in lane, braking to lead vehicle, press a button on the dashboard etc. Importantly, these sensory-motor schemata may also implement active sensing, e.g. “look to the left for cars”. Schemata at higher levels may represent more generic action patterns, situations or task contexts such as “follow the car ahead” or “turn right at an intersection”. The number of levels in this schema hierarchy depends on the modelling purpose (for example, sensory-motor schema may be further divided into basic reflexes and more complex sensory-motor acts).

Schemata are *selected* by virtue of their level of *activation* and the selection process involves dynamical cooperative and competitive interactions between schemata. These interactions are governed by the spreading of bottom-up and/or top-down activation through associative, reciprocal (two-way) links between schemata. These links may be excitatory (positive) or inhibitory (negative) (see Figure 2). The former implement mutual facilitation between schemata “in coalition”, while the inhibitory links implement mutual suppression between competing schemata. The *strength* of the links determines the degree of facilitation and inhibition.

The activation level of a schema is assumed to depend on three main factors: (1) the *total received top-down and bottom-up input activation*, (2) the *lateral inhibition* received from competing schemata and (3) the schema’s *inherent strength*, representing the significance of the schema to the subject. Hence, for example, a strong (salient) input (e.g. a flashing neon sign) is likely to activate a “sign reading” schema bottom-up and inhibit other competing schemata. Moreover, if the sign displays a socially significant word such as ones own name, represented by an inherently strong schema, that schema will be more easily activated than a weaker schema. The strength of schemata as well as the associative links between them is assumed to be determined by associative learning, in line with existing neuroscientific models based on Hebb’s rule [see e.g. [26]).

As a result of this continuous interactive (bottom-up and top-down) flow of facilitation and inhibition, the schemata with strongest activation will win and competing schemata will get suppressed, leading to a temporary stable “attentional” state that (if the selection was successful) reflects the current and anticipated task demands. It should be noted that several schemata, or schema coalitions, may be simultaneously active, as long as there is no competition between them.

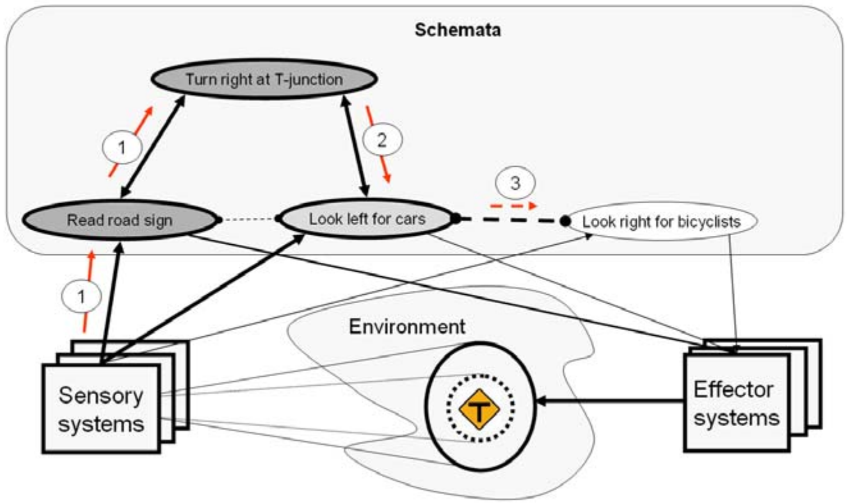

Figure 3. Example of schema selection in the T-junction scenario. Thick lines represent strong associations and thick outline represent inherent schema strength. Shading represents activation level (the darker, the higher activation). See the text for further explanation.

A key feature of this mechanism is that a higher-level schema, representing task context, may initially be activated bottom-up by sensory motor schemata and, in the next step, send context-based top-down activation to bias the competition between underlying sensory-motor schemata. An example of the proposed mechanism is illustrated in Figure 3, in terms of the T-junction scenario illustrated in Figure 1. As the driver enters the T-junction, we assume that he correctly perceives the T-junction road sign. This activates the "standard" "turn-right-at-T-junction" context schema bottom-up (1). As a result, associated sensory-motor schemata such as "look left for cars" are activated top-down, hence implementing "implicit" expectations on which sensory-motor schemata that will be required in the upcoming situation (2). Thus, competing sensory-motor schemata not subsumed by the "standard" T-junction schema, such as "look right for bicyclists", are suppressed since they receive only inhibition and no top-down bias (3). The proposed mechanism thus implies an interactive, context-dependent, flow of bottom-up and top-down activation that results in routine action selection.

A central aspect of this schema selection model is that some schemata (or schema coalitions), are inherently stronger than others (where, as mentioned above, strength is assumed to be determined by associative learning). Strong schemata are easily excited bottom-up and may run automatically, without conscious effort. Thus, schema strength is here proposed as the basis for the classical distinction between *automatic* and *controlled* performance [39]. The suggested mechanism is similar to the concept of neural pathway strength proposed by Cohen, Dunbar and McClelland [5]. Lower-level schemata are assumed to be more easily automatised than higher-level schemata, due to the smaller degrees of freedom involved. However, with sufficient training, even complex tasks may run automatically [29, 39].

If the schema selection mechanism is left to work alone, the strongest schemata will generally win and the only way to override a strong schema is by means of strong bottom-up activation by salient stimuli. However, this is clearly insufficient to deal with novel, complex or inherently difficult situations where the task-relevant schemata may be too weak and/or the task requires a flexible switching between stronger and weaker schemata. This implies the need for a top-down "force" that can bias the schema selection in a certain direction when needed. This is the key role of the supervisory control component described in the next section.

Supervisory control

As described above, if left to operate alone, schema selection is solely determined by the effective input strength and/or the inherent strength of existing schemata. Hence, this will lead to stereotyped behaviour which may be insufficient in novel and/or inherently difficult situations. According to the proposed model, the role of supervisory control, a concept directly adopted from Norman and Shallice [31] is to bias the schema selection mechanism when needed to achieve task goals. This amounts to biasing the underlying competition in favour of potentially weak, but task-relevant, schemata and/or to protect against interference from task-irrelevant distractors. More generally, the role of supervisory top-down bias can be viewed as *stabilising* inherently weak schemata, binding together normally unrelated schemata and overriding inherently stronger schemata when needed. Furthermore, *working memory* can be viewed in terms of supervisory top-down bias with the purpose to sustain activation at the schema selection level in the absence of stimulus input. Since the operation of the supervisory control system involves "pushing" the schema selection system to an inherently non-stable state, it requires the addition of *energy*. Thus, supervisory top-down control is always deployed with *effort* and is assumed to be associated with autonomic physiological responses related to arousal. The present conceptualisation of supervisory control is essentially in line with Kahneman's [17] classical effort-based model of attention.

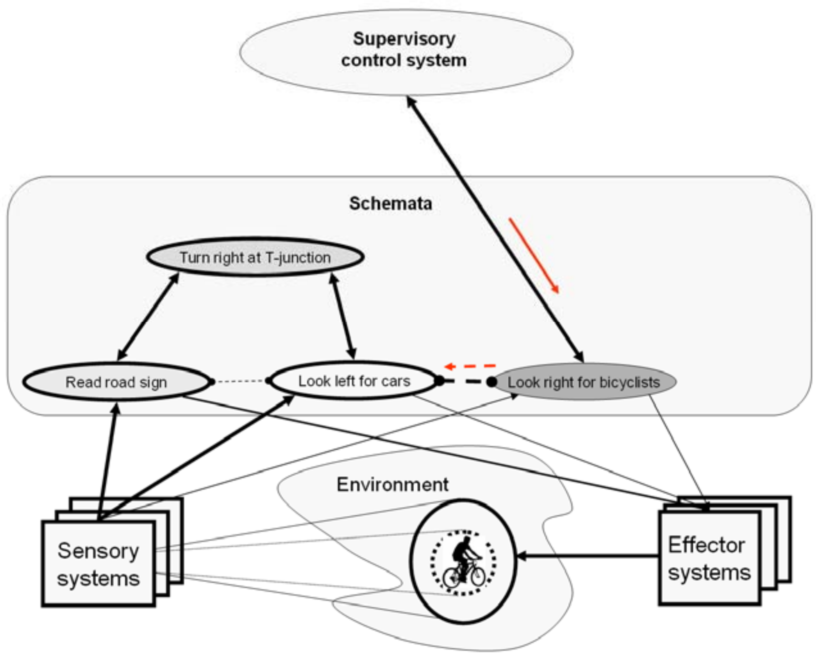

Figure 4 Illustration of how supervisory top-down bias can be used to override the established schema.

The proposed function of the supervisory control component can be illustrated in terms of our T-junction example, as illustrated in Figure 4. Since looking right for bicyclists represent a non-routine action, supervisory top-down bias is needed to override the "look-left-for-cars" sensory-motor schema biased by the "standard" "turn-right-at-T-junction" context schema. If similar situations are frequently encountered, the "look-right-for-bicycles" schema will become increasingly associated with the "turn-right-at-T-junction" context schema until it eventually becomes part of the coalition and receives top-down bias directly from the higher-level schema. Once these associations are established, the supervisory top-down bias is no longer needed and the T-junction negotiation task has again been "automatised". This learning process corresponds to the classical concept of *assimilation* proposed by Piaget [33].

THREE TYPES OF TASK INTERFERENCE

Based on the conceptual model outlined above, three general types of task interference, related to the three model components outlined above, could be distinguished: (1) *interference in sensory- and/or effector systems*, (2) *cross-talk interference between schemata* and (3) *competition for supervisory top-down bias*. This section defines these categories and discusses some further predictions implied by the proposed model.

Interference in sensory- and/or effector systems

This category refers to interference that is solely related to physical sensory/effector systems and not to any internal interference "in the brain". This corresponds to what Wickens [44] refers to as peripheral interference. For example, in the T-junction scenario, delayed or missed responses may occur simply due to the bicyclist appearing in the visual periphery or entirely outside the field of view. This category also includes effector interference such as competition for the hands when peeling a banana while steering.

This mechanism is closely related to change blindness. In our example scenario, the sudden appearance of the bicyclist from behind the tree represents a visual transient, prone to capture attention bottom up. Based on the results from change blindness studies [22, 36], it could be hypothesised that the effect is exacerbated when such a transient onset occurs outside the field of view, since attention capture now has to rely on less salient cues. Some evidence for this effect is offered by a recent study by Lee, Lee and Boyle [22].

Cross-talk interference between schemata

This type of interference occurs due to inhibition between competing schemata and corresponds to the classical notion of structural interference [44]. An example of cross-talk interference at the sensory-motor level is the competition between the "look-left-for-cars" and "look-right-for-bicycles" schemata illustrated in Figure 3 and Figure 4 above. When, as in the case illustrated in Figure 4, supervisory top-down bias is allocated to a sensory-motor schema, competing non-task relevant schemata are suppressed by lateral inhibition. This could be suggested as the basic mechanism behind inattentional blindness [24, 40] as well as looked-but-failed-to-see accidents, where drivers in post-accident interviews report that they did not see the hazard despite looking straight at it [4]. It also accounts for the increased robustness to peripheral distractor stimuli resulting from increased perceptual load [19, 20]. According to the model, these phenomena represent "early selection" in the sense that task-irrelevant schemata are suppressed already at the sensory-motor level. Indeed, this is supported by an EEG-study by Handy et al. [12], which demonstrated reduced neural activity at the lower levels of the visual hierarchy during concurrent performance of perceptually demanding tasks. However, schemata with non-overlapping perception and/or response modalities may still compete at higher levels in the schema hierarchy, which would correspond to the traditional notion of “late selection”.

In the driving context, the model predicts that responses to non-task relevant visual stimuli will be strongly degraded in driving situations with high visual perceptual load, e.g. complex intersections. While this effect has been clearly demonstrated in laboratory studies (e.g. by Lavie and colleagues [19, 20], it has not been isolated in the driving context where it is often confounded with effects related to visual eccentricity (which has led to a great deal of confusion regarding the phenomenon often referred to as “visual tunnelling”).

Competition for supervisory top-down bias

A key implication of the present model is that if two simultaneous tasks both rely on supervisory control bias, performance on one or both tasks will suffer (compared to single task performance) regardless of any cross-talk interference or interference in sensory and/or effector systems. Thus, for example, a working memory task, such a conversation on the phone or with a passenger, may interfere with driving in demanding situations, even if there is no cross-talk interference between schemata at all (all drivers have experienced the need to temporarily interrupt a conversation when entering a difficult driving situation). This can be further illustrated in terms of our T-junction scenario. Recall that our driver was heavily engaged in a phone conversation when approaching the T-junction. As illustrated in Figure 4, supervisory top-down bias is needed to override the dominating “look-left-for-cars” schema to activate the “look-right-for-bicyclists” schema. However, since the phone conversation also demands supervisory top-down bias, there is no top-down bias left for the “look-right-for-bicyclists” schema. Hence, even if the driver “knew” beforehand about the intersecting bicycle lane, he may fail to orient attention towards it due to a *lack of supervisory top-down bias*.

As a consequence of lacking supervisory top-down bias, the schema selection mechanism will operate in bottom-up-driven, automatic, mode where inherently dominant schemata and/or schemata triggered by strongly salient stimuli, will generally win the competition, thus leading to stereotyped, inflexible, behaviour. It would also be expected that, when supervisory top-down bias is allocated to a demanding working memory task, it would be easier for task-irrelevant distractor stimuli to capture attention bottom-up due to the reduced top-down driven inhibition between schemata at the sensory-motor level. Indeed, Lavie and colleagues [20] found increased distractor interference when a working memory task was performed concurrently with a perceptually loading task. Related to this, Strayer et al. [41], in a driving simulator study, found that working memory load did impair post-trip explicit recall of roadside billboards but did not have any effect on the orientation of gaze towards the road signs, a result that is also consistent with the present model (on the assumption that explicit memory storage requires supervisory control but gaze orientation can be automatically triggered bottom up).

The lack of supervisory top-down bias could thus be viewed as another form of “late” selection, where several non-competing lower-level schemata may be activated in parallel, but with a lower level of activation than during single-task performance. In fact, a recent brain imaging study has demonstrated reduced brain activation in parietal association areas during concurrent performance of simulated driving and a working memory task [16].

In the absence of cross-talk interference at lower levels, the model predicts that the interference effect between tasks with non-overlapping sensory-motor demands should be independent of sensory and response modality. This is supported by the large body of work using the Psychological Refractory Period paradigm [e.g. 23, 32, 43]. The notion of lacking top-down bias can also be invoked to explain the finding of Merat and Jamson [25] that the effect of working memory load on a stimulus detection task did not differ when stimulus modality was varied between the visual, auditory and tactile modalities (see [9] for an application of an earlier version of the present model as a theoretical basis for the use of simple detection tasks to assess drivers' selective attention performance).

A further implication of the model is that that working memory load would be expected to interfere with the top-down sequencing of actions in novel or difficult situations. Such impairments are well documented in patients with frontal lobe damage [6], consistent with the widely accepted hypothesis that supervisory top-down control is subserved by frontal brain areas [e.g. 6, 8, 26]. It would be interesting to investigate whether similar impairments may occur in driving as a result of working memory load (e.g. during phone conversation).

DISCUSSION

The present paper outlined an action-oriented view of attention selection in driving with the main purpose to account for proactive, goal- and context-driven selection of actions in everyday driving situations. Mainly based on earlier models by Norman and Shallice [31] and Cooper and Shallice [6], a general attention selection mechanism was proposed where selection occurs through top-down and bottom-up driven processes of cooperation and competition between actions and action patterns represented in terms of schemata. Based on this model, three main types of task interference were suggested and some novel predictions of the model outlined.

The most important difference between the current model and traditional IP models is that the current model rejects the assumption of a strict sequence of information processing stages in favour of a continuous interactive flow of bottom-up and top-down activation in a schema hierarchy, which can be understood and modelled as a dynamical system. As a result, the model naturally captures the active nature of attention selection and accounts for the role of task goals, context and behavioural history, a domain where traditional IP models are notoriously weak.

The model also offers a new perspective on traditional unresolved issues in the IP paradigm, such as early vs. late selection, bottlenecks vs. graded capacity sharing and single vs. multiple resources. In terms of the present model, selection can occur both "early", as a result of cross-talk interference between sensory-motor schemata and "late", as the result of interference between higher-level schema or competition for supervisory top-down bias. However, according to the present model, "low-level" and "high-level" selection would be more appropriate terms. Hence, there is not a single "bottleneck" in the system, but rather many different potential bottlenecks at different levels in the schema hierarchy. With respect to the debate between single and multiple resource theories, the present model postulates multiple "resources" at the schema level, where competition is relatively local and several schemata may potentially run in parallel, while supervisory control can be regarded as a single “resource”, with capacity limited by the total amount of effort that can be mobilised. In this respect, the model is generally consistent with Kahneman [17] who suggested that task interference may occur both due to structural interference between parallel “performance units” (corresponding to schemata in the present model) and an insufficient general supply of effort given the total task demands (corresponding to supervisory top-down bias).

As described above, the model makes several novel predictions that could be tested in future experimental work. In many respects, these predictions differ from those generated by traditional information processing models. In particular, the present model makes it possible to frame clear hypotheses regarding the role of expectancy and adaptive, anticipatory scheduling of attention. Such aspects have been largely ignored in existing experimental work, probably because they are difficult to conceptualise with “standard” information processing models. For example, while existing studies on driver distraction have focused on performance decrements in dual task situations (where subjects are instructed to perform a secondary task while driving), there has been relatively little interest in *why* a driver decides to engage in a secondary task in a certain situation (see [35] for an exception). The conceptual model presented here seems to offer a suitable theoretical framework to address these types of research questions (see Lee, Regan and Young [21] for similar ideas developed within a control theoretical framework).

The proposed action-oriented perspective also has implications for automotive human machine interface design. While, traditionally, a general design goal has been to minimise interference with the driving task (e.g. by reducing the visual demands of displays), the present view also emphasises that, in addition, good HMI designs should support the driver’s adaptive and anticipatory task allocation. Thus, for example, the possibility to interrupt an ongoing task when needed is an aspect of key importance. Moreover, the use of proactive information and performance feedback to support the development of adequate expectations is an area with great potential, especially given recent advances in technologies for vehicle-to-vehicle and vehicle-to-infrastructure communication.

Moreover, the possibilities of measuring brain activation during simulated driving have been exploited in several recent studies [13, 16]. This creates new opportunities to better understand cognitive mechanisms that govern adaptive driver behaviour in general and attention/action selection in particular. The present model, based on an activation dynamics rather than an information processing metaphor, can be naturally linked to this type of data.

Finally, the possibilities of implementing the model computationally are also worth further explorations, as it would enable more specific predictions to be tested in simulation. In this context, a non-linear dynamical systems modelling framework (see e.g. [38]) seems to be particularly well suited, although other approaches, e.g. the computational model developed by Cooper, Shallice and colleagues [6, 7], are worth exploring as well.

REFERENCES

[1] Allport, A. 1987. Selection for action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer & A.F. Sanders (Eds), *Perspectives on perception and action* (pp. 395–419). Hillsdale, NJ: Lawrence ErlbaumAssociates Inc.

[2] Arbib, MA (1995). Schema theory. In MA Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 830-834)

[3] Broadbent, D. 1958. *Perception and communication*. Pergamon Press, NY.

[4] Brown, I.D. 2005. *Review of the 'Looked but Failed to See' Accident Causation Factor*. Road Safety Report No. 60. London: Department of Transport.

[5] Cohen, J.D., Dunbar, K., and McClelland, J.L. 1990. On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, 97, pp. 332-361.

[6] Cooper, R. and Shallice, T. (2000) Contention scheduling and the control of routine activities. *Cognitive Neuropsychology*, 17 (4). pp.297 - 338 .

[7] Cooper, R.F., Schwartz, M.F., Yule, P., Shallice, T. (2005). The simulation of action in complex activities of daily living. *Cognitive Neuropsychology* 22(8), pp.959-1004.

[8] Desimone, R. and Duncan, J. 1995. Neural mechanisms of selective visual attention. *Annual Review of* Neuroscience 18: 193-222.

[9] Engström, J. in press. The Tactile Detection Task as a method for assessing drivers' cognitive load. To appear in G. Rupp (Ed.) SAE special issue on Driver Performance Metrics.

[10] Engström, J. 2008. A model of attention selection in driving. Licentiate Thesis, Chalmers University, Sweden.

[11] Engström, J., & Ljung Aust, M. (in press). Adaptive behaviour in the simulator: Implications for active safety system evaluation. In D. Fisher, M. Rizzo, J. Lee & J. Caird (Eds.),*Handbook of Driving Simulation for Engineering, Medicine and Psychology*: CRC Press.

[12] Handy, T.C., Soltani, M., & Mangun, G. R. (2001). Perceptual load and visuocortical processing: Event-related potentials reveal sensory-level selection. *Psychological Science*, 12, 213-218.

[13] Hsieh L, Young R.A., Bowyer S.M., Moran J.E., Genik R.J., Green C.C., Chiang Y.R., Yu Y.J., Liao C.C., and Seaman S. 2009. Conversation effects on neural mechanisms underlying reaction time to visual events while viewing a driving scene: fMRI analysis and asynchrony model. *Brain research* 1251:162-75

[14] Itti, L. and Koch, C. 2000. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research* 40 (2000) 1489–1506.

[15] Jonides, J. 1981. Voluntary versus automatic control over the mind's eye's movement. In J. B. Long & A. D. Baddeley (Eds.), *Attention & Performance IX* (pp. 187-203). Hillsdale, NJ: Erlbaum.

[16] Just, M.A, Keller, T.A. and Cynkar, J. 2008. A decrease in brain activation associated with driving when listening to someone speak. *Brain Research*, 1205 (2008). p.70-80.

[17] Kahneman, D. 1973. *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

[18] Lamble, D., Laakso, M and Summala, H. 1999. Detection thresholds in car following situations and peripheral vision: implications for positioning of visually demanding in-car displays. *Ergonomics*, Vol 42, No. 6, 807-815

[19] Lavie, N. (2005) Distracted and confused?: selective attention under load. *Trends in Cognitive Sciences, 9,* 75-82

[20] Lavie, N. & De Fockert, J. (2006) Frontal control of attentional capture in visual search. *Visual Cognition*, *14* 863-876.

[21] Lee, J.D., Regan, M.A. and Young, K.L. 2008. What drives distraction? Driver distraction as a breakdown of multi-level control. In Regan, M., Lee, J. and Young, K. (Eds) *Driver Distraction: Theory, Effects and Mitigation*. CRC Press.

[22] Lee, Y.C., Lee, J.D and Boyle, L. 2007. Visual attention in driving: The effects of cognitive load and visual disruption, *Human Factors*, Vol. 49, No. 4.

[23] Levy, J., Pashler, H, & Boer, E (2006). Central interference in driving: Is there any stopping the psychological refractory period? *Psychological Science*,17, 3, 228-235.

[24] Mack, A., & Rock, I. 1998. *Inattentional Blindness.* Cambridge, MA: MIT Press.

[25] Merat, N. and Jamson, A.H. 2008. The effect of stimulus modality on signal detection: Implications for assessing the safety of in-vehicle technology. *Human Factors*, Vol. 50, No. 1, pp. 145–158.

[26] Miller, E.K. and Cohen, J.D. 2001. An integrative theory of prefrontal cortex function. *Annual Reviews of Neuroscience*, 24: 167-202.

[27] Moray, N. 1969. *Attention: Selective processes in vision and hearing*. Hutchison, London.

[28] Moray, N. 1999. Commentary on Goodman, Tijerina, Bents and Wiewille: “Using cellular telephones in vehicles: Safe or unsafe”. *Transportation Human Factors*, 1, 3-42.

[29] Neisser U. 1976. *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco: W. H. Freeman.

[30] 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.

[31] Norman, D., & Shallice, T., 1986. Attention to action: Willed and automatic control of behavior. In R. Davidson, G. Schwartz, & D. Shapiro, (Eds.), *Consciousness and Self Regulation*, Vol. 4, New York: Plenum.

[32] Pashler, H., & Johnston, J.C. (1998). Attentional limitations in dual task performance. In H. Pashler (Ed.), Attention (pp. 155–189). Hove, England: Psychology Press.

[33] Piaget, J. 1971. *Biology and knowledge*. Edinburgh: Edinburgh University Press.

[34] Räsänen, M., & Summala, H. 1998. Attention and expectation problems in bicycle–car collisions: An in-depth study. *Accident Analysis and Prevention, 30*, 657–666.

[35] Rauch, N., Gradenegger, B. and Krüger, H-P. 2008. Identifying user strategies for interaction with in-vehicle information systems while driving in a simulator. In *Proceedings of the European Conference on Human Centred design for Intelligent Transport Systems*, Lyon April-3-4 2008.

[36] Rensink, R. A. 2002. Change Detection. *Annual Review of Psychology, 53*, 245-77.

[37] Shallice,T., Burgess, P.W. 1993. Supervisory control of action and thought selection. in Baddeley,A., Weiskrantz,L. (ed.) *Attention: Selection, Awareness and Control: A Tribute to Donald Broadbent*. Oxford: Clarendon Press, 171-187

[38] Schöner, G. 2007. Dynamical systems approaches to cognition. In: *Cambridge Handbook of Computational Cognitive Modeling*. Cambridge University Press. Ron Sun (editor).

[39] Shiffrin, R.M & Schneider, W. 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127-190.

[40] Simons, D. J., & Chabris, C. F. 1999. Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28*, 1059-1074.

[41] Strayer, D.L and Drews, F.A and Johnston, W.A. 2003. Cell phone induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied*, 9, 23-52.

[42] Summala H, Pasanen E, Räsänen M, Sievänen J. 1996. Bicycle accidents and drivers' visual search at left and right turns. *Accident Analysis and Prevention*, 28(2):147–153.

[43] Welford, A.T. (1952). The ‘‘psychological refractory period’’ and the timing of high speed performance—a review and a theory. *British Journal of Psychology*, 43, 2–19.

[44] Wickens, C.D. 1984. Processing resources and attention. In R. Parasuraman & R. Davies (Eds.), *Varieties of Attention*. New York: Academic Press.

[45] Wickens, C.D. 1992. *Engineering Psychology and Human Performance*, Harper Collins, New York.

[46] Wickens, C.D. 2002. Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3(2), 159-177.

[47] Wickens, C.D. Letter to the editor. Transportation Human Factors, 1, 1999. 205-206.

[48] Wickens C.D. and Horrey, W.J. 2009. Models of attention, distraction, and hazard avoidance. In M.A Regan, J.D. Lee and K.L. Young (Eds.) *Driver Distraction: Theory, Effects and Mitigation*. Boca Raton: CRC Press.