Visuomotor Control, Eye Movements, and Steering: A Unified Approach for Incorporating Feedback, Feedforward, and Internal Models

Otto Lappi
Cognitive Science
Department of Digital Humanities
University of Helsinki.

Callum Mole
School of Psychology
University of Leeds.

Psychological Bulletin
© 2018 American Psychological Association
0033-2909/18/$12.00
http://dx.doi.org/10.1037/bul0000150

Visuomotor Control, Eye Movements, and Steering: A Unified Approach for Incorporating Feedback, Feedforward, and Internal Models

Otto Lappi
University of Helsinki

Callum Mole
University of Leeds

The authors present an approach to the coordination of eye movements and locomotion in naturalistic steering tasks. It is based on recent empirical research, in particular, on driver eye movements, that poses challenges for existing accounts of how we visually steer a course. They first analyze how the ideas of feedback and feedforward processes and internal models are treated in control theoretical steering models within vision science and engineering, which share an underlying architecture but have historically developed in very separate ways. The authors then show how these traditions can be naturally (re)integrated with each other and with contemporary neuroscience, to better understand the skill and gaze strategies involved. They then propose a conceptual model that (a) gives a unified account to the coordination of gaze and steering control, (b) incorporates higher-level path planning, and (c) draws on the literature on paired forward and inverse models in predictive control. Although each of these (a–c) has been considered before (also in the context of driving), integrating them into a single framework and the authors' multiple waypoint identification hypothesis within that framework are novel. The proposed hypothesis is relevant to all forms of visually guided locomotion.

Public Significance Statement
There is a gap between current understanding of human driving performance in experimental psychology, engineering, and neuroscience. This review traces the common historical and conceptual roots, identifies major open issues and puts forward proposals for future research. The work has implications for any domain involving skilled movement through the world, from fundamental experiment design and theory to modeling of real-world performance, through to designing autonomous vehicle/driver assistance technologies.

Keywords: visuomotor control, forward models, waypoints, eye movements, steering

In the control of high-speed locomotion, humans display a remarkable capacity to encode complex visual information and transform it into well-coordinated motor commands. We rely on this ability in everyday life when cycling to work in the morning, running along a forest path or steering a car through a series of bends. Most forms of sport are based on it. How appropriate complex actions are executed with such speed, precision, and apparent ease is one of the outstanding unresolved questions in understanding how our brains' visual, motor, and spatial navigation systems work together and can support such a wide range of skilled behaviors.

An excellent example of these behaviors—ubiquitous in modern society[1]—is driving. Driving is a complex real-word task subserved by many fundamental perceptual-motor and attentional functions. For example, it has been argued that driving well exemplifies (among others) selective visual attention (Wolfe & Horowitz, 2004), computation of salience and task relevance (Tanner & Itti, 2017), target interception and obstacle avoidance (Regan & Gray, 2000), and the coordination of eye movements in naturalistic settings (Kowler, 2011).

Otto Lappi, Cognitive Science, Department of Digital Humanities, University of Helsinki; Callum Mole, School of Psychology, University of Leeds.

Preparation of the manuscript was made possible by mobility supported by an Erasmus+ agreement between the University of Helsinki and the University of Leeds, the Academy of Finland project Mulsimco (279905, OL), the Experimental Psychology Society (Callum Mole), the University of Leeds Research Mobility Programme (Callum Mole) and the LS2N École Centrale de Nantes (Otto Lappi). Discussions with Richard M. Wilkie, Gustav Markkula, Esko Lehtonen and Jami Pekkanen have been instrumental in developing the ideas. The manuscript was finalized and submitted whilst the authors were visiting researchers at École Central de Nantes, hosted by Franck Mars.

Correspondence concerning this article should be addressed to Otto Lappi, Cognitive Science, Department of Digital Humanities, University of Helsinki, P.O. Box 9, 00014, Finland. E-mail: otto.lappi@helsinki.fi

[1] World Health Organization statistics give a number of 1.78 billion vehicles registered for road use at the end of 2015 (http://apps.who.int/gho/data/node.main.A995; retrieved 2017/03/04). This does not include vehicles in military or sports use.

1

The final article is available at:
http:// doi.org./10.1037/bul0000150

Abstract

The authors present an approach to the coordination of eye movements and locomotion in naturalistic steering tasks. It is based on recent empirical research, in particular, in driver eye movements, that poses challenges for existing accounts of how we visually steer a course. They first analyze how the ideas of feedback and feedforward processes and internal models are treated in control theoretical steering models within vision science and engineering, which share an underlying architecture but have historically developed in very separate ways. The authors then show how these traditions can be naturally (re)integrated with each other and with contemporary neuroscience, to better understand the skill and gaze strategies involved. They then propose a conceptual model that (a) gives a unified account to the coordination of gaze and steering control, (b) incorporates higher-level path planning, and (c) draws on the literature on paired forward and inverse models in predictive control. Although each of these (a– c) has been considered before (also in the context of driving), integrating them into a single framework and the authors' multiple waypoint identification hypothesis within that framework are novel. The proposed hypothesis is relevant to all forms of visually guided locomotion.

Keywords: visuomotor control; forward models; waypoints; eye movements; steering

Public significance statement: There is a gap between current understanding of human driving performance in experimental psychology, engineering, and neuroscience. This review traces the common historical and conceptual roots, identifies major open issues and puts forward proposals for future research. The work has implications for any domain involving skilled movement through the world, from fundamental experiment design and theory to modeling of real-world performance, through to designing autonomous vehicle / driver assistance technologies..

1. Introduction

In the control of high-speed locomotion humans display a remarkable capacity to encode complex visual information and transform it into well-coordinated motor commands. We rely on this ability in everyday life when cycling to work in the morning, running along a forest path or steering a car through a series of bends. Most forms of sport are based on it. How appropriate complex actions are executed with such speed, precision and apparent ease is one of the outstanding unresolved questions in understanding how our brains' visual, motor and spatial navigation systems work together and can support such a wide range of skilled behaviours.

An excellent example of these behaviors - ubiquitous in modern society[1] - is driving. Driving is a complex real-word task subserved by many fundamental perceptual-motor and attentional functions. For example, it has been argued that driving well exemplifies (amongst others) selective visual attention (Wolfe & Horowitz, 2004), computation of salience and task relevance (Tanner & Itti, 2017), target interception and obstacle avoidance (Regan & Gray, 2000) and the coordination of eye movements in naturalistic settings (Kowler, 2011).

However, much of our current understanding of visuomotor performance is extrapolated from experiments that aim to characterize very specific aspects of behavior, by experimentally isolating stimulus features and subtask elements. Consequently, there is often a fairly large gap between the proposed individual mechanisms and being able to actually (quantitatively) model their contributions in concrete domains which the understanding gleaned from experiment is intended to generalize to (e.g. driving). Furthermore, when multiple mechanisms are recruited by a complex task, they could interact in ways which cannot be predicted from models of the mechanisms themselves. Therefore, although many proposed mechanisms are *intended* to be task-general one is hard pressed to find models that can account - in quantitative detail – for complete task performance in representative natural tasks, where the core processes are orchestrated for skilled performance.

What is needed is that core cognitive processes be integrated into unified models that can (i) be expressed in rigorous quantitative terms, e.g. computer simulation and (ii) account for full task performance, rather than a specific aspect of the task, in an ecologically realistic task setting (Newell, 1973). We firmly believe that human perception, action and cognition can only be understood through such integrated models of full performance in tasks of complexity and demand representative of the tasks people actually engage in their real everyday and professional life.

In modelling driving, fortunately, solid progress has been made on quantitative theories which attempt to explain precisely how visual information can be used to support high-speed locomotion. Driving, therefore, represents a useful test-bed of a complex real-world task subserved by multiple perceptual, cognitive and motor control mechanisms: a domain of skill which can be analyzed in detail, yet understood in a principled way in order to generalize to other domains of visuomotor skill, eye-hand coordination and visual control of locomotion.

Perhaps the most generalizable subtask of driving is using vision to steer a path in 3D space (lateral control). The capacity to steer a path on the basis of visual preview information and correct for positional errors is a basic tenet of any locomotor task, from crawling, and walking, to piloting aircraft and land vehicles. Over the years, there have been many explicit and computationally implemented steering models of the visual and motor control strategies humans may employ when they operate a motor vehicle, making it one of the best understood complex real-world visual tasks (for reviews, see McRuer, 1980; Land, 1998; Wann & Land, 2000; Macadam, 2003; Lappi, 2014).

Interestingly, these models broadly emerge from two traditions which tend not to reference one another: on one hand there is *vision science and experimental psychology*, and on the other hand there is *vehicle dynamics and control engineering.* (The same control theoretical core ideas are also present in modern computational movement neuroscience, but we will defer discussion of this connection till later, once we have first introduced the general ideas in a concrete domain).

In the vision science literature, driving is analyzed as control based on perceptually available optical information (Gibson, 1958; Land & Lee, 1994; Boer, 1996; Land, 1998; Kim & Turvey, 1999; Wann & Swapp, 2000; Wilkie et al., 2003; Salvucci & Gray, 2004; Sentouh et al., 2009; Saleh et al., 2011; Boer, 2016). This research is grounded on rigorous experimental work, and builds on what is known about the functional limitations of the human perceptual system. A weakness of the vision science models is that many are underspecified or at any rate not computationally implemented (e.g. Wann & Swapp, 2000). Those that are computationally explicit (e.g. Salvucci & Gray, 2004; Boer, 2016), on the other hand, tend to be developed to explain human performance on simplified laboratory tasks (which may lack the stimulus complexity of the real world), rather than driver performance in real world tasks. They therefore are quite limited in the scope of steering behaviour they can account for (i.e. they may perform well on slow-speed lateral control tasks such as lane keeping or lane changing, but fall short of being able to capture higher levels of skill (Macadam, 2001) or non-routine behaviours such as near-collision manoeuvres, Markkula, 2014)).

Conversely, the vehicle dynamics and control literature provides computational models - based on control-theoretic and vehicle dynamics concepts and used to reproduce driver behavior accurately for driver-in-the-loop simulations - which do describe many aspects of driver behaviour well, including complex paths and on-limit-handling (Sharp, Casanova & Symonds, 2000; Macadam, 2001; Prokop, 2001; Sharp & Valtetsiotis, 2001; Sharp, 2005; Keen & Cole, 2006; Cole, Pick & Odhams, 2006; Odhams & Cole, 2009; Keen & Cole, 2011, 2012; Timings & Cole, 2012). A weakness of *these* models is that they, in turn, incorporate little about the function and limitations of human physiology, and hardly anything about what is known of the perceptual and cognitive system of the human driver. This is an acknowledged shortcoming (Nash, Cole & Bigler, 2016).

The engineering models and psychological models share the same basic architecture, as they originate from the same body of work done in the 1970s (generally the last papers to be referenced in both literatures are McRuer et al., 1977 and Donges, 1978). But for the last 40 years these two bodies of literature have developed somewhat independently of one another.

To very briefly outline the history, we note that the engineering work developed from feedback models of manual control in aircraft pilots in the 1950's and 1960's (McRuer & Krendel, 1956, 1959; McRuer et al., 1965; Kleinman, Baron & Levison, 1971) and driver models developed through the 1970's (Wierville, Gagne & Knight, 1967, McRuer et al., 1975; Allen & McRuer, 1979). These are reviewed in McRuer & Jex (1967) and McRuer (1980), respectively. More recently the engineering models have incorporated computational ideas borrowed from the analysis of complex control problems in robotics and artificial intelligence (on estimation and prediction, motor planning, and optimization in trading off accuracy against effort, reviewed in Macadam, 2003; Sharp & Peng, 2011 present a technical overview of techniques and applications). Such models therefore are much more computationally sophisticated, and complex, than the steering models of vision science. The vision science models tend to favor stimulus-response models or control laws where steering is based directly on perceptual variables in the stimulus, in large part due to the ecological traditions of Gibson (1958) and Lee (1976). That is, behavior is accounted for as responses to available optical information, avoiding assuming complex internal representations or computations in the brain (Warren, 2007; Lee et al., 2009; Zhao & Warren, 2015).

We believe that there is great potential for the theoretical integration of the approaches. On the one hand, incorporating what is known about human physiology, behavior and cognition into the engineering would allow these practical models more accurate representations of the actual mechanisms underlying steering skill. But on the other hand, the strict requirement to *account for full task performance in extended sequences of behavior*, the advances in computational techniques already rigorously applied to this real-world domain in the engineering literature would be equally advantageous for developing more unified models in the psychological literature, offering a real-world test bed for testing and developing such models beyond individual mechanisms accounting for specially constructed laboratory tasks. Indeed, we believe that the best hope for understanding the visual control of steering lies in combining such complementary viewpoints with one another - and we envision that the extensive literature on computational movement neuroscience provides a natural framework for such conceptual integration.

In what follows we briefly outline the origin and state of the art of steering models in vision science. We then show the most promising points of integration, in terms of both underlying concepts and accounting for recent experimental work. We end with presenting an integrative framework that (1) takes a unified approach to treating coordination of gaze and steering as the primary control task, (2) incorporates an account of higher-level path planning and (3) draws on the literature on forward and inverse internal models in predictive control from theoretical neuroscience.

Although each of (1-3) has been considered in before (also in the context of driving), integrating them into a single framework and our *multiple waypoint identification hypothesis* within that framework are novel. While we discuss the problem in the concrete and well-understood case of driving, nothing about the general ideas depends on special aspects of the domain, therefore our arguments have direct relevance to other, less well understood, forms of visually guided locomotion. Our approach is not an applied one: we do not ask how knowledge gleaned from "basic" laboratory tasks is to be "applied" in a specific domain. Our interest is, instead, in what a detailed understanding of this behavior can contribute to the fundamental understanding of skilled visual control of locomotion.

2. Compensatory and Anticipatory Control - a Brief History of 40 Years

The papers by McRuer et al. (1977) and Donges (1978) represent a fairly discrete point where the two traditions leading to current steering models in engineering and vision science diverged[2]. Drawing from pilot and driver models of the 60s and 70s, these seminal papers described steering control as consisting of nested hierarchies of *navigation*, *guidance* and *control* (**Figure 1A**), with control itself being further partitioned into sub-levels; there were three in the case of McRuer et al. (**Figure 1B**), two in the case of Donges. Interestingly, while Donges (1978) is the most heavily referenced 70's steering model paper in the vision science literature (via Land, 1998 and Salvucci & Gray, 2004), this paper is actually much less prominent in the engineering literature which cites McRuer et al. (1977) and its later developments.

The bottom level of control, *compensatory* control, is envisioned as closed-loop feedback control maintaining lane position by cancelling observed lane position error relative to some reference value (desired lane position provided from outside the feedback loop). However, it has been well known for decades that inertia, momentum, perceptual-motor delays, and driver-vehicle system delays mean that during high-speed steering error develops too quickly for compensatory control to be sufficient. A higher level of *anticipatory* control (called *pursuit* in McRuer et al., 1977), is therefore needed. This level of control *anticipates* upcoming changes in required heading, in order to keep error within acceptable limits. This higher-level control is not based on lane position error only, but depends on "preview information" from the road scene further ahead. The preview information is assumed to be available through vision and - in some way or another - to specify the *desired future path*. The process of extracting this relevant information from the road scene was called guidance, which is situated between control and navigation in the overall wayfinding task hierarchy (**Figure 1A**).

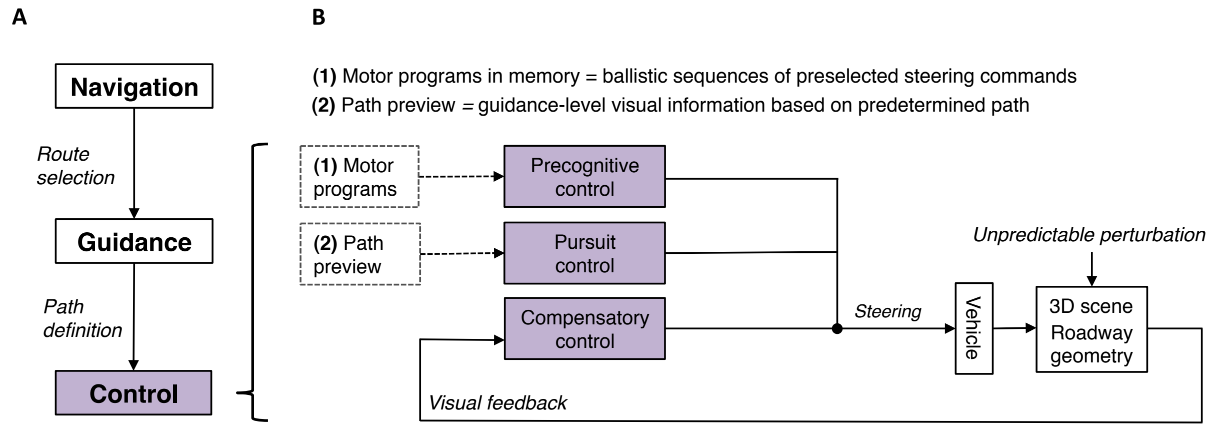

Figure 1. The McRuer et al. (1977) framework. This framework (minus precognitive control) was also the basis of Donges' (1978) "two-level" framework. **A** The place of visual control of steering within the hierarchical organization of wayfinding. **B** The McRuer et al. (1977) control level organization aligned with compensatory ("feedback") and pursuit/precognitive ("feedforward") control.

In both papers (McRuer et al., 1977; Donges, 1978), anticipatory control is often called "open-loop", or "feedforward" control, as it receives as input guidance information – that is, road or desired future path curvature at different distances - as opposed to only the difference between actual and desired lane position, which is the case for *compensatory* control.

McRuer et al. (1977) also posited a third level of "precognitive" anticipatory control based on "learned patterned open-loop responses" (ibid., p394). Precognitive control was conceptualized in terms of the *motor program* idea, according to which very fast and accurate overlearned motor skills are based on a ballistic read-out from long term memory, because (due to inherent physiological delays in feedback) they are too rapid for feedback control (Keele, 1968; Schmidt, 1975; see also Schmidt, 2003; Shea & Wulf, 2005; Summers & Anson, 2009). During preparation of a motor program the order, timing and magnitude of the movements in a sequence would be determined before initiation of the first action in the sequence, and once retrieved from memory and initiated the program should run its course without feedback.

Precognitive control was not included in Donges' (1978) two level model, and is consequently rarely discussed in vision science models (e.g. Land, 1998; Salvucci & Gray, 2004) or traffic psychology literature (though see Godthelp, 1985; Wallis et al., 2007), which historically follow more on Donges' (re)formulation than the McRuer et al. directly.

The 1970s engineering models were introduced to the vision science literature in the 1990s by Michael Land and colleagues (Land & Lee, 1994; Land & Horwood, 1995; Land, 1998; see also Salvucci & Gray, 2004; **Figure 2**). Both McRuer and Donges were vague on precisely what preview information underpinned anticipatory control or how it was supposed to be extracted from the scene. In vision science, however, there was already a long history of searching for optical information – such as *optic flow* (Gibson, 1958; Lee & Kalmus, 1980; Koenderink, 1986), *retinal flow* (Regan & Beverly, 1982; Kim & Turvey, 1999), *tau* (Lee, 1976) or *affordances* (Fajen, 2013) – that could serve as steering cues[3], and for control laws linking specific cues to steering response (Wann & Swapp, 2000; Fajen, 2001; Fajen & Warren, 2003; Warren, 2006; for reviews see Regan & Gray, 2000, and Wann & Land, 2000).

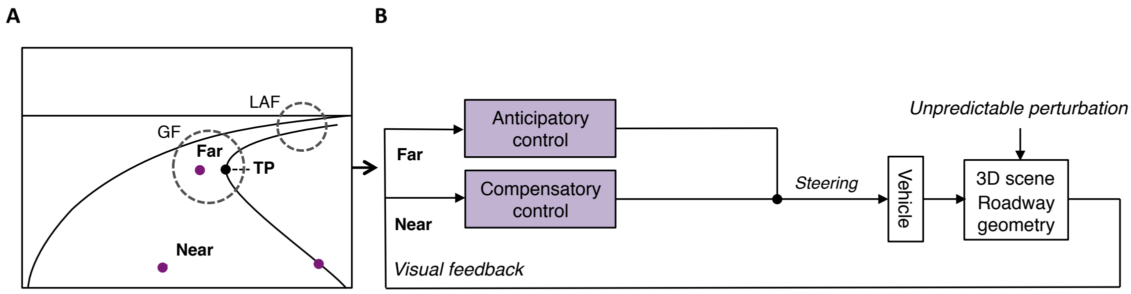

Figure 2. The two-point control framework (Land, 1998; Salvucci & Gray, 2004), loosely based on Donges (1978). **A.** Schematic representation of the visual field indicating the location of the Near and Far zones when steering into a bend with four putative steering points from the literature. A near point on the road centre or road edge are monitored peripherally. **B** Near and far steering point visual direction and rate of movement feedback are used for compensatory and anticipatory (prospective) steering control, respectively. Note that while gaze control is not explicitly modelled, this framework is in part inspired by observations of the ubiquitous visual strategy of Guiding fixations (GF) into the far region - either the tangent point (TP; Land & Lee, 1994; Land, 1998) or a travel point on the future path (Salvucci & Gray, 2004; Boer, 2016). Look-ahead fixations (LAF) (Lehtonen et al., 2013; 2014) intermittently scanning the road further ahead with "gaze polling" saccades (Wilkie et al., 2008) go beyond the two-point framework, however.

The advantage of the vision science approach is that theoretical and model development has been informed by a large body of relevant experimental work on laboratory analog steering tasks that can manipulate the availability and quality of visual feedback (e.g. Land & Horwood, 1995; Hildreth et al., 2000; Fajen, 2001; Wilkie & Wann, 2002; 2003; 2006, Fajen & Warren, 2003; Wallis, Chatziastros & Bülthoff, 2003; Wallis et al., 2007; Mars, 2008; Wilson et al., 2007, 2008; Cloete & Wallis, 2009; Li & Cheng, 2011, 2013; Mars & Navarro, 2012; Frissen & Mars, 2014; Kountouriotis et al., 2016; Mole et al., 2016), as well as more naturalistic field research on the gaze strategies drivers actually use to sample task-relevant information (e.g. Shinar, McDowell & Rockwell, 1977; Land, 1992; Land & Lee, 1994; Land & Tatler, 2001; Chattington et al., 2007; Lappi et al., 2013; Lehtonen et al., 2013, 2014).

In a seminal contribution to this research tradition, Land & Horwood (1995) empirically linked the two bottom levels of steering control (as suggested by McRuer and Donges) with specific portions of the visual scene. They demonstrated that visibility of a simulated road scene could be restricted to 1 degree (vertical) near and far segments, supposedly without loss in steering accuracy whilst steering a 'tortuous' road. Steering using far road information only would be smooth but imprecise, whereas steering using near road information would be jerky, but accurate. These properties led Land (1998) to propose far road information and near road information feed into anticipatory and compensatory control, respectively (**Figure 2AB**)[4].

In another influential paper that addressed the question of where drivers look when steering through a series of bends, Land & Lee (1994) showed that much of the time the driver's gaze was directed towards the apex region of the upcoming bend: at or very near the lane edge *tangent point* (**TP**; **Figure 2A**). TP is the point on the edgeline where the visual orientation of the projection of the edgeline in the driver's visual field is reversed; looking at the TP will make the line of sight tangential to the lane edge. This "TP orientation" strategy, as it became to be called, was incorporated within the two-level framework when Land (1998) posited a two-point control framework with the TP as the *far point*, and a *near point* on the lane edge next to the vehicle as steering points.

Salvucci & Gray (2004) implemented a two-point model where steering output is determined as a weighted combination of the visual direction of two steering points, a far point and a near point, which were modelled as moving along with the observer at a fixed distance ahead (i.e. the modelled steering points were *travel points* on the *future path*). This Proportional-Integral feedback control model drew on the experimental psychology literature on combining multiple cues to perceive heading relative to a target or to determine the steering required to intercept a target (Fajen 2001; Wann & Wilkie 2004). It has inspired a number of similar implementations (Sentouh et al., 2009; Boer, 2016) and has been incorporated into the ACT-R cognitive architecture, used to investigate multitasking while driving (Salvucci et al., 2005; Salvucci, 2006).

Largely on the basis of Land & Lee (1994) and Salvucci & Gray (2004), a two-point framework (with the tangent point as the far point) was quickly established as the predominant approach to understanding the visual control of locomotion in curve driving (e.g. Underwood et al., 1999; Marple-Horvat et al., 2005; Chattington et al., 2007; Coutton-Jean et al., 2007; Wilson et al., 2007; Authié & Mestre, 2012; see also Boer, 1996; Hildreth et al., 2000), and the default textbook account of how we steer and where we look when we do so.

2.1. Prospective vs. Predictive Anticipation

Comparing **Figure 1** and **Figure 2** shows how the Land (1998) and Salvucci & Gray (2004) two point models formalised a conceptual shift away from the open-loop, feedforward conception of anticipatory control described in McRuer et al. (1977). This is a subtle shift and may go unnoticed if – as is not uncommon – the terms “anticipatory”, “feedforward” and “open-loop” are used interchangeably. But whereas Donges stated categorically that “an anticipatory reaction is always feedforward” (1978, p.698), in the two-point models anticipatory control and compensatory control are *both* seen as closed loop processes, using information from a far point (through foveal vision) and a near point (through peripheral vision), respectively.

This is a significant historical development and the point where the current psychological and engineering literature on the visual control of locomotion most differ. We will in the next section review some more recent empirical developments that seriously challenge the simplified view of the two-point architecture. But before that we must clarify precisely what the conceptual differences in different treatments of anticipatory control are.

The engineering models traditionally rely heavily on *state estimation* and *prediction* (i.e. internal models), and current state-of-the art engineering models represent the driver with *model-based predictive optimal control* models (see below). The vision science models (**Figure 2**), however, are not *predictive* controllers but rely on what Zhao & Warren (2015) call *prospective* control. Prospective control is “anticipatory” control, but simply a response to *the present value of a predictive variable* (such as a far point, model-free control), not a response to the *predicted value of a variable* (from a forward model of the environment, model-based control). It is thus a *non-representational* account of anticipation that does not require an internal model of the world (scene layout) or predictive capability.

It is very important to differentiate between predictive, prospective and ballistic forms of anticipatory control. To avoid conflating these three very different concepts about anticipatory control, and to make more clear why it leads to nontrivial ambiguity, we propose some terminological conventions in **Table 1**. The key difference is that predictive control is model-based, whereas prospective and ballistic control are not.

Table 1

Classification of compensatory and different forms of anticipatory control

Type of control	**Compensatory**	**Anticipatory**		
		Prospective	**Precognitive**	
			Ballistic	**Predictive**
Source of feedback	Local stabilizing information	Visual preview of road ahead	Visual preview of road ahead	Visual preview of road ahead
Response generation	Direct responses to unpredictable perturbations	Direct responses to "preview" cues. (*currently* observable optical variables correlated with *future* state).	"Motor programs" (neuromuscular code, stored in long-term memory) Not sensitive to feedback information about outcomes once launched	Based on both forward internal models (predictive, sensory/spatial code) and inverse models (motor code). May incorporate feedback and efference copy even after action is initiated
Examples	Earliest controller models without pursuit control / preview	Classical steering models in vision science, such as the two-point models	McRuer et al. framework; 1980s traffic psychology models	Model-based optimal control driver models in vehicle dynamics

Online, bottom-up, feedback models, ***continuous*** *responses to constantly available sensory feedback*

Feedworward models; sampling of feedback and/or generation of actions can be ***intermittent***

Note. Local stabilizing information here refers to near road visual information and vestibular and kinesthetic feedback.

Calling compensatory control feedback and anticipatory control feedforward can be particularly misleading because prospective control *is* feedback control. Predictive control, on the other hand is "precognitive" like a motor program (top-down, feedforward, open loop) in the sense that it can run offline, in the absence of sensory feedback. However, whilst both are classified under "precognitive" control in **Table 1**, an important distinction can be made between predictive control and the original ballistic motor programs. Ballistic control is a purely open-loop sequence of *motor commands* and does not incorporate feedback until complete (Keele, 1968; Schmidt, 1975; McRuer et al., 1977, Donges, 1978). Predictive control, on the other hand, is based on *internal forward models* that estimate the current and/or the future state of the world – including likely consequences of actions – and *internal inverse models* that select the optimal action, given current goal and world state estimated by the forward model.

Forward models integrate information across multiple sensory sources and use an efference copy of outgoing motor commands to predict sensory consequences, to which the subsequent feedback can then be compared. The feedback is then used to update forward models in learning a better internal model of the world (including the controlled system), so that prediction error is minimized. The value of this more complex type of architecture (**Figure 3AB**) is that the perceptual-motor system can respond to the (best estimate of) the current state of the world rather than sensory information which may be incomplete, noisy and laggy due to inherent physiological delays. For motor planning, forward models allow the system to produce and evaluate a prediction of the likely consequences of actions.

Internal model based prediction is a mature computational approach and the dominant account of predictive control in engineering (e.g. Cole, Pick & Odhams, 2006; Keen & Cole, 2011; Boer 2016) and in computational movement neuroscience (e.g. Miall & Wolpert, 1996; Wolpert & Kawato, 1998; Wolpert, & Ghahramani, 2000; Desmurget & Grafton, 2000; Wolpert, Diedrichsen, & Flanagan, 2011). It is yet to be incorporated into the vision science research on steering.

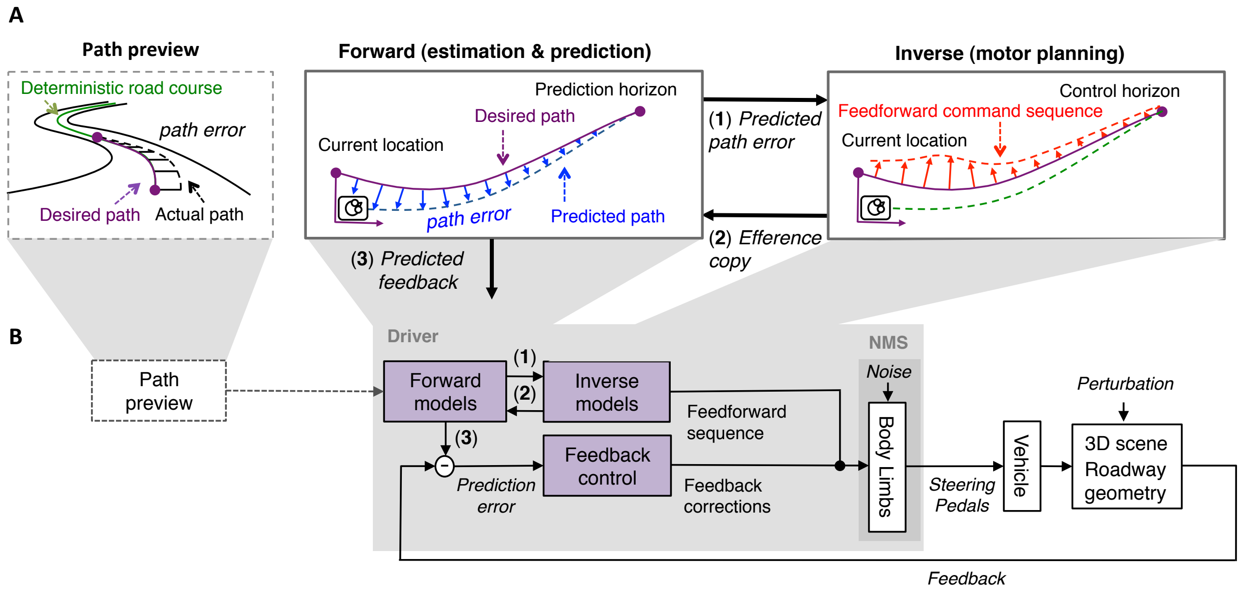

Figure 3 A,B General architecture representing the state-of-the-art of implemented control theoretical driver models in vehicle system dynamics engineering (cf. Prokop, 2001; Hoult & Cole, 2008; Keen & Cole, 2006, 2011; Nash, Cole & Bigler, 2016). Here, "anticipatory" control is treated as truly feedforward. Predictive internal forward models determine where current state and steering command lead to, and the future path error (distances to points on the desired path) is passed to the inverse feedforward controller which computes the optimal sequence of motor commands. Residual error due to unpredictable disturbances drives the feedback controller. Paired forward/inverse models and feedback/feedforward control levels means a motor plan does not unfold completely "ballistically" from a prespecified motor program, but from a filtered state estimate. NMS = neuromuscular system.

To summarise, in the 40 years since the "split" between the vision science and engineering literatures, the direction vehicle dynamics models have thus taken quite different direction from the developments in vision science. In rather stark contrast to the two-point reinterpretation of the two-level model (**Figure 2**) which is a simple proportional feedback control, the engineering models are based on a feedback/ feedforward architecture (**Figures 3A & 3B**) as new and more sophisticated modelling methods have been adopted founded on powerful computational ideas of internal models and predictive optimal control, and the feedforward level is based on optimal control ("planning") and internal models ("observers" in modern control theoretical terminology).

Model based prediction is a more sophisticated solution to unstable feedback control (i.e. due to feedback delays) than the purely ballistic motor programs envisioned in the 1970s engineering models. In the next section we suggest how such control based on internal models may be better able to account for some observed high-level behaviours – in particular, intermittency and look-ahead fixations – than the current state of the art.

3. Internal models and predictive control in driving

Naturally, it depends on one's broader theoretical outlook whether one considers the move to treating anticipatory control as model-free prospective control a backwards step holding back model development or a desirable development towards theoretical parsimony. Likewise, the internal model-based (predictive) framework can be either considered a powerful tool that allows us to more adequately explain skilled human behavior, or an overly complicated and elaborate way to account for behaviors that can be explained more simply.

One of the reasons for the shift towards prospective models in vision science over the past 40 years may be the strong historical predilection for maximally *parsimonious* models that, as far as possible, derive control of action *directly* from optical variables in the stimulus without intervening computations (cf. Gibson 1958; Warren, 2007; Zhao & Warren, 2015). After reviewing the experimental literature on locomotor and interceptive actions, Zhao & Warren (ibid.) argue that there is ample experimental evidence that can be modelled by prospective (online) control laws based on currently available visual variables, and no convincing experiments to date where the data demand that one posit complex internal models (a similar argument is made by proponents of radical "embodied cognition", e.g. Wilson & Golonka, 2013).

We suspect that another, methodological, reason may be restricting model domains to highly controlled laboratory analog steering tasks (simple flat curve geometry, linear control dynamics, often sparse visual input presented as a 2D projection on a screen, and trials lasting only a few seconds). When the task itself is very simple, then it may not be a surprise that a simple steering model using steering points with short time headway can capture human trajectories reasonably well (e.g. Wilkie & Wann, 2003; Salvucci & Gray, 2004; Wilkie et al., 2008; Boer, 2016). But such tasks may not *discriminate* well between model-based and model-free control. So, it is perhaps no coincidence that model-based control is used in engineering driver-in-the-loop simulation models that attempt to capture highly *complex* skill (i.e. driving involving high levels of expertise, such as race driving, Keen & Cole, 2011) - whereas the prospective two-point model framework is still the predominant modelling approach for *simple* lane keeping and curve negotiation tasks in experimental psychology (Wann & Land, 2000; Lappi, 2014).

While pure feedback two-point models with an additive combination of stabilizing and guidance level information from egocentric visual direction of steering points *can* successfully perform simple steering manoeuvers in a simulation (Boer, 2016), they can be criticized on a number of empirical and theoretical grounds. For example, some perceptual inputs such as flow speed (Kountouriotis et al., 2016), are not easily incorporated into the two point model in its current form (Mole et al., 2016). In addition, it has been long recognized that a real driver does not expend continuous attention to the visual control of steering, but samples from the scene *intermittently* (Senders, 1968; Blaauw, Godthelp & Milgram, 1984; Johnson et al., 2014; Kujala et al., 2016). Also field observations of gaze strategies in natural driving have revealed that the hypothesis of a single "steering point" in the far zone is not sufficient to account for the complexity of fixation behavior in rich natural contexts (Lappi, Lehtonen, Pekkanen & Itkonen, 2013; Lehtonen et al., 2013, 2014; Lappi, 2014; Lappi, Rinkkala & Pekkanen, 2017). We believe that it is in these last two examples, intermittency and gaze strategies, where integrating engineering and vision science principles could lead to significant progress, and the potential role of internal models can be best assessed.

3.1. Intermittency of Perception and Action

One aspect of visual control of steering that online feedback approaches do not handle well is intermittency in visual sampling and control (Blaauw, Godthelp & Milgram, 1984; Godthelp, Milgram & Baauw 1985; Godthelp, 1985, 1986; Hildreth et al., 2000; Johnson et al., 2014; Markkula, 2014; Johns & Cole, 2015; Markkula et al., 2017; see also Senders, 1968; Land & Furneaux, 1997).

Pure feedback control models are designed for continuous feedback, and to produce continuous and incremental adjustment[5]. They work when the relevant state variables (e.g. steering points) are assumed to be continuously observable and when the dynamics of the controlled system are sufficiently low-frequency with respect to latencies in the feedback loops.

Although naturalistic eye tracking studies (e.g. Land & Lee, 1994; Lappi, Rinkkala & Pekkanen, 2017) have shown that drivers do spend a high proportion of time looking to the "far" region where guidance-level information is presumed to come from in the two-level framework (Donges, 1978; Land, 1998; Salvucci & Gray, 2004; Frissen & Mars, 2014; Mole et al., 2016), it has been long acknowledged in the literature (Senders et al., 1968; Godthelp, Milgram & Blaauw, 1984) that in the real world humans do not allocate their undivided attention to the road ahead. Looking at landmarks, traffic signs and instruments, potential hazards, other road users, or in-car communication and infotainment devices all break the connection between where the driver is looking and where they are going (as do *look-ahead fixations*, to which we return in more detail in the next section). This results in intermittency in visual sampling - which is an essential characteristic of eye movement strategy while driving, but a difficulty for models based only on compensatory control and prospective anticipation (i.e. pure online feedback control).

It remains unclear how the perceptual-motor system deals with intermittency, and whether or what roles internal models play in it. As outlined in the historical sketch above, early papers proposed that skilled handling of such events was possible through feedforward ballistic control (McRuer et al., 1977; see also Godthelp, 1985). The driver was envisioned as opening and closing control loops based on some supervisory control strategy (though the mechanisms involved in the control decisions were somewhat unclear, and outside the scope of formalization in the models). Specifically, ballistic motor programs would handle intermittency by 'taking over' during periods when visual feedback is unavailable. Such accounts would lead the driver to switch between feedback and feedforward control (without the need for internal models in the control-level model).

More recently, explicit modelling of intermittent perception and action has become a hot topic in both vision science and engineering. A distinction can be made here between *intermittent visual sampling* (e.g. Senders et al., 1968), and *intermittent motor control* (e.g. Markkula, 2014). The former can be characterized as "intermittent observation, continuous decisions" and the latter as "continuous observation, intermittent decisions" - whereas prospective control is "continuous observation, continuous action", while, of course, real actions could even be "intermittent observation, intermittent action".

In the driving context (as in a number of laboratory contexts, for review see Bosco et al., 2015) intermittent sampling can be experimentally examined using visual occlusion paradigms where periodically the participant's view is blocked, allowing the effect of removing visual feedback on performance to be examined (Senders et al., 1968; Milgram, Godthelp & Blaauw, 1982; Godthelp, 1985; Cavallo et al., 1988; Wallis, Chatziastros & Bülthoff, 2002; Wallis et al., 2007; Macuga et al., 2007; Cloete & Wallis, 2009; Johnson et al., 2014; Kujala et al., 2016; Johns & Cole, 2015; Pekkanen et al., 2017). In the vision science steering model literature, there are as yet no formal treatments of intermittency but at a metaphorical level a number of authors have suggested that an "image"

(Senders et al., 1968) or a "visual buffer" (Land & Furneaux, 1997; see also Tatler & Land, 2011) may retain sufficient visuospatial information to guide steering.

In other words, such accounts assume that when visual preview and feedback are unavailable the driver maintains adequate performance by updating an *internal model* of the relevant control variables' values, such as distance to lead car, or time to line crossing. To avoid immediate obsolescence, the internal model would need generative predictive capability, and could also be *updated* on basis of (non-visual) sensory feedback (see also Loomis, Klatzky & Giudice, 2013; Lappi, 2016; cf. "expectancy" in Näätänen & Summala, 1974, 1976). The scene could then be sampled again when uncertainty increases to unacceptable levels (Senders et al., 1968; Godthelp & Blaauw, 1982, Johnson et al., 2014; Kujala et al., 2016) or when a variable is estimated to have reached a threshold value (Godthelp, Milgram & Blaauw, 1984).

Intermittent control can also be approached from the motor end: modelling steering as intermittent bursts of (ballistic) control with continuous sampling or continually available perceptual information available from a forward model or "buffer" (Markkula, 2014; 2017; cf. Johns & Cole, 2015, and Gawthrop et al., 2011). Steering control is treated a sequence of feedforward open-loop adjustments based on internal models that estimate and predict the state variables based on an efference copies of the steering movement (Franklin & Wolpert, 2011). An adjustment is triggered when evidence accumulates that an ongoing action is not consistent with current goals (Markkula, 2014: Markkula et al., 2018; see also Gawthrop et al., 2011). The advantages of such an approach are that a) it allows for complex and adaptive control during occlusion (Johns & Cole, 2015) and b) unlike the Senders et al. (1967), Blaauw, Godthelp and Milgram (1984) and Sullivan & Ballard (2014) models it does not relate internal model uncertainty to active sampling (e.g. in the Markkula (2014) model the accumulation of *evidence* leads to *steering action*, rather than accumulation of *uncertainty* leading to *sampling*).

There is considerable evidence that human drivers can easily handle periods of intermittency of visual feedback and still successfully perform both real and simulated lane keeping (Blaauw, Godthelp & Milgram, 1984; Hildreth et al., 2000; Johns & Cole, 2015), lane changing (Godthelp, 1985; Macuga et al., 2007) and curve negotiation (Godthelp, 1986; Cavallo et al., 1988; Kujala et al., 2016). A natural explanation for how the perceptual-motor manages to cope with intermittency is that drivers use internal models. However, this view has also been challenged on the grounds that under (severe) restrictions to available visual feedback it appears steering control is not as robust as one might assume on basis of "precognitive" control mechanisms (ballistic or predictive).

Specifically, Wallis and colleagues (Wallis et al., 2002; 2007) demonstrated that even the simple and presumably highly overlearned and automatic manoeuvre of a lane change was surprisingly dependent on visual feedback. In real-world occlusion studies, the road often can only be occluded for more than a few seconds, and feedback is generally available at the end of the trial (though see Macuga et al., 2007). To avoid intermittent feedback or knowledge of results allowing participants to shift strategies, Wallis et al., (2002, 2007) used a simulator where all visual feedback (during or after the maneuver) was removed. Under these conditions only the initial wheel-turn (that takes the vehicle to the other lane but leaves it pointing diagonally) was executed - the second wheel turn (an equal turn in the opposite direction which aligns the vehicle with the lane) was not. The executed manoeuver was strikingly similar to turning a corner, and appeared to show that participants held a severe "misconception" about the relationship between steering wheel movements and the resulting path of the vehicle – despite many years of driving experience.

Restoring performance to adequate levels in the Wallis et al. tasks did not require *continuous* feedback, though. It was sufficient to provide (very) short intermittent glances providing they were appropriately placed at the end of the first phase of the sequence.

Such behaviour would seem to suggest that humans do not after all adequately perform *sequential* motor actions formed of multiple discrete actions “precognitively” (cf. Godhtelp, 1985; Hildreth et al., 2000). This is a challenge for the motor program idea, but also model predictive control which suggest humans should be able to maintain a forward model up to some reasonably long prediction horizon, and inversely plan a feedforward sequence of motor outputs up to some control horizon (**Figure 3A**, for complex maneouvers this is typically in the order of 4s, e.g. Johns & Cole, 2015).

Wallis et al. (2002, 2007) themselves propose a “turn and look” strategy – that only the first (intermittent, ballistic) motor action in a sequence is planned and initiated top-down, but the parameter setting and initiation of the follow-up actions require appropriately timed bottom-up feedback from the sensory consequences of the first action (i.e. a discrete visual sample)[6]. Wallis et al. (2002; 2007) experiments thus suggest that *intermittent feedback is needed* even in simple and highly familiar sequential tasks. Under the “turn and look” strategy it is not necessary to assume that drivers possess a multiphase internal model in order to explain behavior in this specific task. If indeed they do, the results imply that such models’ prediction and planning horizons are surprisingly limited. Rather than a detailed representation of complex path and a motor plan to achieve that path, the default mode for drivers would be to plan only a short way ahead, perhaps up to the level of the current *guiding fixation* (see discussion of guiding and look-ahead fixations in the next section)[6].

What one cannot infer from Wallis et al. (2002, 2007), though, is that human drivers *cannot* form more detailed sequential plans. Indeed, it appears that evidence for the “turn and look” strategy might only be obtained in an artificial trial-based setting where knowledge of results is denied (whereas the driving task is an extended behavioral sequence). The authors themselves point out that great care was taken to prevent the participants “becoming aware” that their performance in the previous trial was below par, to prevent a "strategy" shift or a "learning effect". A task-general unified model of a human driver should of course be able to perform such a strategy shift. It also remains an open question as to which strategies - look and turn or predictive - humans would spontaneously use under conditions with richer stimuli and, possibly, more demanding maneuvers requiring higher levels of situational awareness.

In any case, the Wallis et al. results do highlight how the timely control of visual *sampling* is an essential component of the steering strategy. If the default strategy employed in real driving is indeed akin to “turn and look”, then appropriate timing of the next look must be programmed as part of each phase in the sequence (else the sequence may break). The next section examines the co-ordination of eye-movements and steering in more detail.

3.2. Guiding fixations and look-ahead fixations

Even in the case where sampling intermittency is not due to visual demands of targets competing with the driving task (e.g. monitoring hazards, distraction) there appears to be intermittency inherent in the gaze strategy subserving skilled actions. Specifically, a stereotypical interleaving of guiding fixations (GFs) and look-ahead fixations (LAFs) can be observed in a wide variety of sequential natural tasks (Land and Furneaux 1997; Land, Mennie & Rusted, 1999; Pelz & Canosa, 2001; Hayhoe et al., 2003), as well as in the laboratory (Ballard, Hayhoe & Pelz, 1995; Mennie, Hayhoe & Sullivan, 2007).

GFs serve to guide ongoing action “just in time” (Ballard, Hayhoe & Pelz, 1995; Ballard et al., 1997; Land & Hayhoe, 2001) while LAFs appear to anticipate later actions, with gaze returning to guidance after the LAF (Pelz and Canosa 2001; Mennie, Hayhoe & Sullivan, 2007)[5]. The ubiquity of this gaze strategy suggests that sampling intermittency is not simply something the perceptual-motor system needs to 'handle', but could potentially have a fundamental functional role in coordinating skilled action.

In driving, such sampling behaviours manifests in interleaving GF fixations to the "far" road with LAFs on the road even further ahead (**Figure 2A**; cf. also **Figure 4A** in the next section**)**. This *gaze polling* (Wilkie et al., 2008) visual strategy is a robust and obvious characteristic of natural eye movements in driving (see visualization in Lappi, Rinkkala & Pekkanen, 2017)[7].

For GFs there is a close coupling between gaze and steering, with GFs typically preceding steering with a lead time about 1s (Land, 1992; Land & Tatler, 1999; Land, 2007; Chattington et al., 2008; Lehtonen et al., 2014). For LAFs this cross correlation is absent and LAFs are typically directed much further ahead than the ~2s time headway which is characteristic of GF and considered useful for *online* control (Boer, 2016). These *spatial* and *temporal* differences appear robust across a number of naturalistic studies (Lappi & Lehtonen, 2013; Lehtonen et al., 2013, 2014). Note that there is a major difference between *lead time* and *time headway*, although the typical numerical values are similar. Lead time is the delay from gaze rotation to steering wheel rotation and hence *a priori* independent of speed and where one looks. Time headway on the other hand is the time to reach the current *point of fixation* at the current *speed* when one fixates a point on one's future path.

It is maybe worthwhile to point that we have been careful to define GFs & LAFs in terms of their spatiotemporal properties, not their presumed *functional* properties. The operational concept itself contains no specific assumptions about the exact targeted points for GFs and LAFs. In typical bends, the GF region in particular contains a number of putative targets postulated by different accounts of steering control, such as the tangent point (Land & Lee, 1994, cf. Land, 1998), edges of "the acceptable trajectory envelope" (Mars & Navarro, 2012), travel points on the future path referenced to the tangent point (Boer, 1996; Mars et al., 2011; Mars & Chevrel, 2017; Lappi, Lehtonen, Pekkanen & Itkonen, 2013), travel points at about 2 second time headway ahead (Lehtonen, Lappi, Koirikivi & Summala, 2014; cf. Salvucci & Gray, 2004) and/or waypoints at various distances (Wilkie, Wann & Allison, 2008; Lappi, 2014). Here the distinction is used independently of any particular steering model or modelling approach. GFs do attractively align with "guidance-level" control in two level framework, whereas the role of LAFs is clearly more anticipatory, but these should be considered functional hypotheses not a definition of the GF/LAF distinction itself. Importantly, in the basic two-level framework there seems to be no natural functional role for LAFs. The exact implications of gaze polling (GF-LAF-GF-LAF) patterns for steering control therefore remains an important open question for the development for models of the visual control of locomotion.

Guiding fixations are coupled to the ongoing subtask of positional control, leading action by ~1s and therefore target locations where an operation is about to be performed (cf. far steering points - travel points or waypoints - as in e.g. Land & Lee, 1994; Salvucci & Gray, 2004; Wilkie, Wann & Allison, 2008; for reviews see Land, 1998; Land, 2006; Lappi, 2014) or objects about to be manipulated imminently (Land & Hayhoe, 2001). Gaze polling (LAFs) are used to glance "ahead" to targets of later actions (Ballard, Hayhoe & Pelz, 1995; Pelz & Canosa, 2001; Mennie, Hayhoe & Sullivan (2007)[8], and so in locomotor tasks, could have a role in *trajectory planning* (Wilkie, Wann & Allison, 2008; Lehtonen et al., 2013, 2014): a higher level of steering control compared to online compensatory and prospective guidance control of two-point models where visual feedback is immediately translated into steering response.

When the scene and path geometry are simple, steering models using steering points with short time headway can fit human trajectories well (e.g. Wilkie & Wann, 2003; Salvucci & Gray, 2004; Wilkie et al., 2008; Boer, 2016). Also, the spatiotemporal characteristics of GFs appear to be in line with the "far" steering points posited by such accounts (Land & Lee, 1994), challenging the need for model-based control. However, when path geometry becomes more complex and longer previews are available model-free control runs aground. Availability of preview of a scene (path) systematically lead to LAF behaviour, which breaks the tight coupling between gaze and steering and are typically targeted too far to make a *direct* contribution to steering response, therefore remain problematic for any 'online' control model of steering. LAFs may allow skilled drivers to take advantage of the preview: for example, when steering a chicane a skilled driver might alter their approach to the initial bend to allow a smooth approach to the second bend (an example of this behaviour was observed when multiple sequential slalom gates were visible in Wilkie et al., 2008).

Model-free, 'online', prospective control strategies do not predict that drivers would alter the current unfolding steering movement in anticipation of an upcoming steering command. But if one assumes the existence of internal models (posited by a number of models in movement neuroscience and engineering) principled functional interpretations of the role of LAFs become possible, and we can begin to see a way to state locomotion and eye movement coordination as a unified control task that can be stated in a control theoretical framework. We propose that a potential mechanism for LAFs could be to *extend the planning horizon from the current guiding fixation waypoint to the look-ahead fixation landing point* as an added waypoint (looking where you want to go). On the motor side, LAFs may *facilitate programming of a sequence of lateral and longitudinal control actions that will lead to intercepting the gaze landing points* (going where you look). In other words, the hypothesis is this: the point of fixation of a LAF is made into a new *waypoint* in the technical sense that a motor plan is generated to intercept it, *via* other waypoints, including the GF point of fixation and previously fixated waypoints.

We propose that this *multiple waypoint identification hypothesis* gives a very natural interpretation of LAFs in the visual control of steering, including but not restricted to the domain of driving. On the other hand, it is hard to see how it could be stated within the framework of prospective model-free control. Within the model predictive control framework it becomes feasible that LAFs serve selection and/or parameter setting of motor plans (i.e. inverse models). And a complementary role for LAFs may be maintaining and updating an internal representation of visual space needed to support such complex, multi-step plans (i.e. forward models; or, cognitive functions functions that in the general vision science literature are variously called "visual buffer", "transsaccadic memory" or "spatial image"). These conceptual possibilities will be further explored in the final subsection.

3.3. Neural underpinnings of steering

Eye movement behavior (and steering performance) can be accurately recorded in naturalistic settings, providing a useful means to investigate active visuomotor strategies. But the gaze data reviewed above does not alone suffice to determine the underlying control mechanisms (such as that proposed in the previous section).

Brain imaging work on different levels of wayfinding behaviors (from determining current lane position and heading, to steering to an optically-available egocentric goal, to planning a route from memory in allocentric space, **Figure 1A**) has begun to reveal the neural underpinnings of the processes hypothesized in the steering models. Many of the processes recruited in driving have been studied in restricted form, and neural substrates have been identified for many of the essential (non-steering) component subtasks, such as self-motion/egocentric heading estimation (Dukelow et al., 2001; Peuskens et al., 2001; Huk, Dougherty & Heeger, 2002; Morrone et al., 2000; Smith et al., 2006; Wall, & Smith, 2008; Cardin & Smith, 2009; Pitzalis et al., 2009; Smith, Wall & Thilo, 2012; Pitzalis et al., 2013), visuospatial object tracking and short term memory (Wolbers, 2008; Nummenmaa et al., 2016), oculomotor planning and execution (Pierrot-Desseilligny et al., 2003, 2004; Munoz, 2002; Girard and Berthoz, 2005; Krauzlis, 2004), and wayfinding (Burgess, Maguire & O'Keefe, 2002; Epstein & Vass, 2014; Barry & Burgess, 2014; Meister & Buffalo, 2016; Epstein et al., 2017), to name a few. Discussing the neural underpinnings of any of these would be a topic for a review of its own and beyond present scope. We will instead focus on the fMRI work that is most *directly* relevant to the two-level framework as presented above (and its planned extension according to the multiple waypoint identification hypothesis), centering on: (i) the neural basis of visuomotor control of steering (Field, Wilkie & Wann, 2007; Billington et al., 2010; Billington, Wilkie & Wann, 2013), and (ii) simulated car driving (Walter et al., 2001 Li et al., Spiers & Maguire, 2006, 2007ab; 2012 Kan et al,. 2013; Schweitzwer et al., 2013).

The laboratory work on the neural basis of steering uses active locomotor control (steering) tasks with path preview. These tasks and associated stimuli are specifically designed on the basis of the two-level steering framework to separate the compensatory and anticipatory processes. But the interpretations heavily rely also on the literature on estimation of heading from optic flow patterns (Peuskens et al., 2001), (predictive) target tracking (Wolbers et al., 2008) and internal forward and inverse models in perception and motor planning (in the context of visually guided hand and eye movements) (Blakemore et al., 2001; Miall, Reckess & Imamizu, 2001; Ramnani, 2004; Fernandez-Ruiz et al., 2007). We will return to this rather important tension between the design and interpretations in the final two sections.

On the other hand, some brain imaging work has used immersive virtual reality scenarios to represent richer environments with more complex driving tasks, such as navigating 3D scene layouts (e.g. intersections), and interacting with moving objects (e.g. traffic). Naturally, this means that the stimulus and task are much less controlled. Therefore, the contrasts become more open to interpretation, and the exact nature of the component processes underlying the participants' performance is not yet well-understood, i.e. the performance is explained more in terms of common sense and analogy from individual laboratory tasks, rather than an integrative model of performance at the complex task. (In section 4 we discuss what such a model would look like, on the basis of current knowledge in psychology and neuroscience).

The primary body of work into the neural basis of steering has been conducted by Wann & colleagues (Field, Wilkie, & Wann, 2007; Billington, Field, Wilkie & Wann, 2010; Billington, Wilkie & Wann, 2013, but see also Huang, Chen & Sereno, 2015). In these studies subjects travel across a flat textured ground plane which provides an optic flow stimulus to elicit a sense of self-motion and changing heading. The task is to actively steer a simple course specified by complete road-edges (Field et al., 2007) or slalom gates (Billington et al., 2013), or "passively steer" (so that all participants received the same visual stimulation but no feedback on accuracy) a course specified by road-edges (Billington et al., 2010). Preview course information (visibility of "far" road edges) can be manipulated to produce contrasts that reveal areas involved in guidance in the two-level framework sense.

Across the three studies it seems that activation of the posterior parietal cortex (PPC), posterior parts of the superior and inferior parietal lobules, and extending medially into the precuneus region was associated specifically with preview information. Other robust areas of activation appear to be the cerebellum (associated with skilled active motor control and internal forward models), and MT+ (with MT classically considered sensitive to optical motion and MTS for optically and extraretinally specified self-motion), and the PPC (superior parietal lobule, SPL, inferior parieta lobule, IPL or precuneus; classically associated with visuomotor coordinate transformations, visuospatial attention and short-term memory).

Billington et al. (2013) suggest that key functionally important areas of the PPC are the SPL and the precuneus. It has been suggested that these areas a role in maintaining target information during delays/occlusions (Fernandez-Ruizet al., 2007; Wolbers et al., 2008), which led Billington et al. (2013) to posit that in the steering tasks the SPL and precuneus have a role in representing "intended goals in visual coordinates" and "predicting the future location of objects in order to execute timely movements" (Ibid. p. 2). The conceptual backdrop for Billington et al.'s (2013; see also Field et al., 2007; Billington et al., 2010) discussions was waypoint tracking and path planning, which (as we have discussed) does not neatly align with two-level models (although the distinction is subtle, it is critical). In our proposed multiple waypoint hypotheses, however, these functional proposals potentially correspond to identifiying waypoints in visual co-ordinates (SPL), and holding multiple waypoints in short-term memory (precuneus). (See section 4 for further discussion).

The neural basis of higher level processes in the wayfinding hierarchy (guidance & navigation, see Figure 1A) has been investigated with fMRI in a number of simulated driving papers (Walter et al., 2001; Li et al., 2012; Kan et al., 2013; Schweitzwer et al., 2013; reviewed in Lappi, 2015), including a series of studies on expert taxi drivers by Spiers & Maguire (Spiers & Maguire, 2006, 2007ab). Taxi drivers (and controls) were asked to navigate around the streets of a virtual city of London. Afterwards, they were interviewed by showing them replays of their own drives and asked to recollect what they were looking at, thinking and doing at each moment in time. The verbal protocols were analyzed qualitatively to identify discrete episodes (such as coasting, route planning, turning, expectation confirmation and violation etc.). The epochs corresponding to these episodes were used as the contrasts for fMRI activation. Episodes involving steering control (such as turning) produced similar activation to laboratory steering control tasks, i.e. in the frontal premotor areas and posteriorly in parieto-occipital areas. A number of other studies on more complex (and unconstrained) simulated driving have also consistently identified dorsal occipital/temporal and PPC activation (Walter et al., 2001; Li et al., 2012; Kan et al., 2013; Schweitzwer et al., 2013).

However, there was also activation in the prefrontal cortex (PFC) and the retrosplenial cortex (RSC), not observed in the steering studies probing near/far contrasts, and only rarely observed in simulated driving studies with simpler tasks . Intriguingly, RSC connects the parahippocampal and occipital "place areas" associated with encoding viewpoint-specific place information and the entorhinal cortex and hippocampal complex (associated with navigation and spatial memory, Epstein et al, 2017), and with the PPC, which is associated with coordinate transformations for visuomotor control and maintaining target information across saccades and occlusions and oculomotor programming (Crawford, Henriques & Medendorp, 2011). RSC, almost invariably involved in spatial memory tasks (Mitchell et al., 2017), has been functionally linked with "piecing together" route and scene information from locally observed snapshots (Park & Chun, 2009; Vann, Aggleton & Maguire, 2009) or "anchoring" observed scene information to cognitive maps in long term memory (Epstein et al., 2017).

To summarize, in tasks constrained to steering (i.e. not navigation or making driving decisions) the most consistent activations (**Figure 4)** are seen in (i) PPC, especially in the SPL and intraparietal sulcus (IPS), probably often involving the functionally identified parietal eye field (PEF), (ii) the occipito-parietal cortex extending medially into the precuneus region, and (iii) the human occipitotemporal motion complex (MT+), which comprises of functionally identified areas middle temporal (MT) which responds to visual motion, and middle superior temporal (MST) which responds more strongly to coherent flow consistent with self-motion, as well as vestibular stimulation (Morrone et al., 2000; Dukelow et al., 2001; Huk et al., 2002; Wall & Smith, 2008; Smith, Wall & Thilo, 2011). Active motor tasks engage areas in the precentral gyrus (primary motor area M1) and premotor cortex (PMC) in the middle and superior frontal gyrus, which also contains the functionally identified frontal eye fields (FEF), as well as the cerebellum (**Figure 4**, solid boxes). As can be expected, tasks involving complex scenes and higher "level" processes in the wayfinding hierarchy involve a more widespread activation of brain areas, including areas classically associated with long-term memory and planning. Activation has been observed in the posterior cingulate/retrosplenial cortex (PCC/RSC), parahippocampal (PHC) and entorhinal/hippocampal complex (HC), as well as prefrontal cortex

(PFC), especially dorsolateral (dlPFC) and ventrolateral (vlPFC) surfaces and the frontopolar cortex (aPFC), sometimes also ventromedial, orbital and insular cortex (**Figure 4**, dashed boxes).

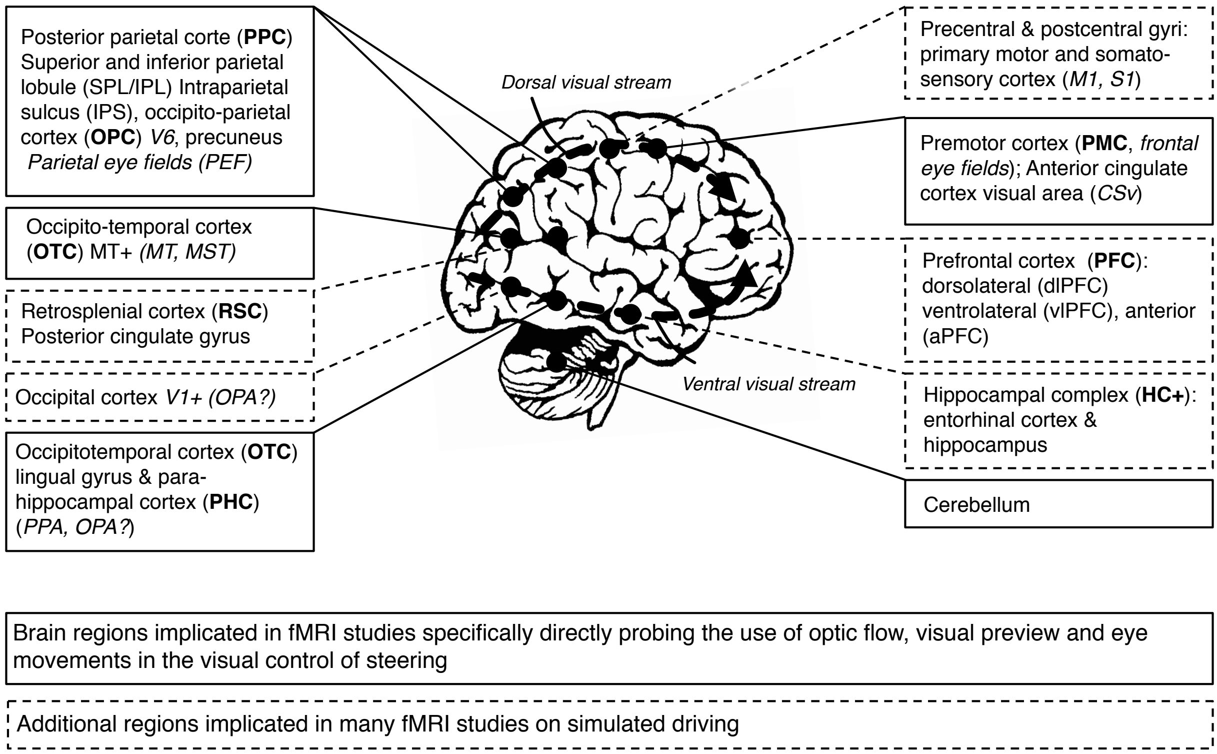

Figure 4. Schematic representation of brain areas activated in fMRI experiments involving visuomotor steering tasks and in simulated driving tasks. PPC = Posterior parietal cortex; SPL = Superior parietal lobule; IPS = Intraparietal Sulcus; *PEF* = Parietal eye field; *FEF* = Frontal eye field; *CSv* = Cingulate gyrus visual area; *MT+* = Human visual motion complex (*V5/MT* & *MST*); RSC = Retrosplenial cortex; aPFC = Anterior prefrontal cortex; vlPFC = Ventrolateral prefrontal cortex dlPFC = Dorsolateral prefrontal cortex; PHC = Parahippocampal cortex, *PPA* = Parahippocampal place area, *OPA* = Occipital place area; HC = Hippocampus; M1 = Primary motor cortex; PMC = Premotor cortex, *FEF* = Frontal eye field.

4. Precognitive Control: A Modern Approach

Even within the experimental vision science literature, traditionally with a penchant for parsimonious control law models, there is a recognized need to expand beyond the assumption of steering as an additive combination of stabilizing and guidance level information (Frissen & Mars, 2014; Mole et al., 2016). Furthermore, observational and experimental studies of gaze strategies have rebutted the notion of guidance level information being recovered from (looking at) a single “steering point” (Lappi, 2014; Lappi, Rinkkala & Pekkanen, 2017). If sampling and control needs to be modelled as fairly discrete, intermittent processes, a more powerful architecture than prospective control is required. What could such an architecture look like?

Figure 5 presents a framework extending the prevailing “two-level” account of how humans steer, which integrates ideas of model-based predictive control from engineering (e.g. Keen & Cole, 2006; Odhams & Cole, 2009, 2010; Keen & Cole, 2011) and computational neuroscience (for reviews see Wolpert, & Ghahramani, 2000; Wolpert, Diedrichsen, & Flanagan, 2011). By placing the co-ordination of gaze and steering at center stage and allowing for predictive control at all levels of the wayfinding hierarchy, the expanded framework takes into account both the sequential (intermittent) character of steering (Section 3.1) and gaze (Section 3.2).

This explicit emphasis on the coordination of gaze and steering through GF/LAF is a novel concept, and, we submit, a necessary step for future development of steering models. In particular, there are three desiderata that are at the moment not fulfilled in the visual steering models literature, but which the proposed architecture could provide. First, the sophisticated gaze strategies observed in visual sampling in naturalistic tasks (such as gaze polling between guiding and look-ahead fixations) need to be addressed in a principled way (**Figure 5A**). Second, models should be consistent with and draw on the relevant neuroscience (**Figure 5B**). Finally, there should be a natural way to express the models as minimal extensions of existing control theoretical frameworks (**Figure 5C**) in order to be able to build on the large body of work in vision science and engineering. We will next go through each of these desiderata in turn.

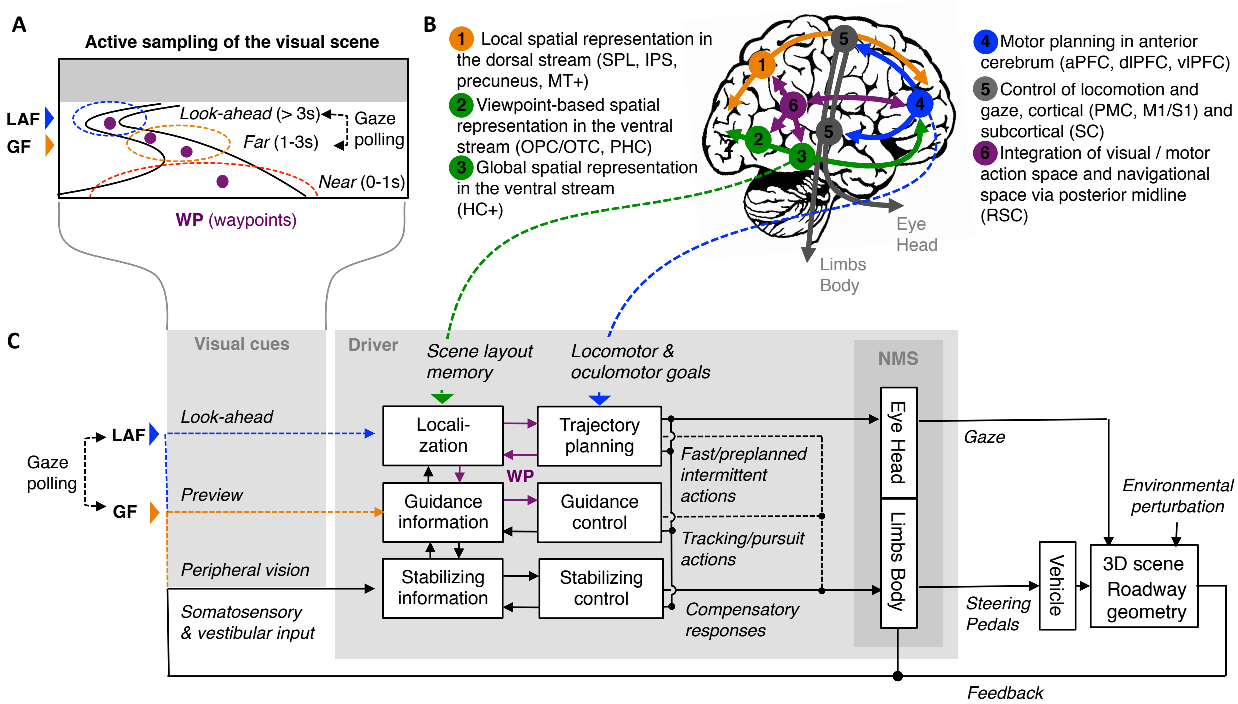

Figure 5. Framework for a unified visuomotor model of the control of steering and gaze. **A** Waypoints identified and tracked by guiding fixations (GF) and look ahead fixations (LAF). **B** A sketch for a plausible neurological basis of the framework, drawing from general cognitive neuroscience and the emerging body of work on the brain imaging of visual control of steering and simulated driving. **C** Three-level framework with coordination of gaze and steering as the core control problem, and paired inverse and forward models at all levels of hierarchy. The aim is to account for how intermittent visual sampling serves action coordination by structuring input into the brain, and also aids in sequencing motor commands (see text for explanation). This gives a more concrete interpretation of the multiple-waypoint hypothesis that accounts for the functional roles for GFs & LAFs as designating locomotor goals (section 3.2.), and the observed patterns of brain activity (section 3.3.). NMS = neuromuscular system; SC= Superior colliculus; for abbreviations of other brain regions please see Figure 4.

Driving is an example of sophisticated visuomotor coordination with robust and highly repeatable gaze strategies, so it seems reasonable – indeed essential - to treat driving as a single unified eye-hand coordination task (i.e. oculomotor and locomotor planning and control) rather than treating gaze strategies, steering strategies attention and memory etc. as separate problems (cf. Land & Furneaux, 1997; Lappi, 2016). A more complete model of visuomotor mechanisms underlying driving, therefore, will not only account for lateral and longitudinal steering, but also how the timing and direction of visual sampling behaviour is coordinated with steering control.

Although a close coupling between gaze and steering is qualitatively consistent with two-point models in that many theories explicitly relate moment-to-moment changes in gaze to steering outputs (Land & Lee, 1994; Wilkie et al., 2008; Salvucci & Gray, 2004), these models only deal with the sensory inputs available *given* that the driver is viewing a specified target in the road scene. The coordination of gaze and steering is in other words only implicitly assumed, and not explicitly part of the formal steering model.

But if it is assumed a driver will look at a far steering point (Land & Lee, 1994) and perhaps monitors a near steering point peripherally (Land, 1998; Salvucci & Gray, 2004), then clearly some process in the brain must specify the gaze target, and control gaze so that is maintained on the desired point. Treating gaze and steering as independently modeled subtasks crucially leaves out the question of how gaze and locomotion are coordinated: the steering model simply assumes that “gaze control” maintains gaze on the target, and gaze control does not need to account for the generation of steering movements. We claim that treating gaze and steering as different aspects of a *unified* task is required to allow the visual sampling intermittency caused by LAFs and GFs to be understood.

In particular, we (Section 3.2.) and others (Wilkie et al., 2008; Lappi, 2014) have proposed that the driver identifies *waypoints* at various distances in the scene and steers to intercept them. We then interpreted the functional roles for different aspects of gaze behaviour as follows: GFs and LAFs track current and specify new waypoints, respectively. In order for this hypothesis about GF and LAF to be modelled, we propose it is necessary to bring back a *third* level - McRuer’s precognitive control, only in a modern predictive control such a third level would not amount to simply performing pre-learned, habitual, motor patterns.

In **Figure 5C** we refer to the third level as *Trajectory planning* and *Localization*. Crucial to this conception is the paired inverse-forward models idea and the existence of a (intermittent) feedback loop connecting the models via the world. Inverse processes would determine a sequence of steering inputs *given* a path to follow (in visual preview), whilst the waypoint specification idea states that eye movements can actively shape the path.

The interpretation is that through directing gaze further ahead in space (LAFs) the trajectory planning system can *specify* the desired path for the coupled forward model (*Localization* in **Figure 5C**). This would be an indexical strategy (called “deictic” processing Ballard, 1997 or “cross cueing” in Gazzaniga, 2013); that is, a strategy depending on the fact that the inverse (trajectory planning) and forward (localization) systems are connected not only directly, but through the world: motor commands shift gaze which changes the retinal image which acts as new input. In effect, rather than having to fully specify the desired path (via efference copy, in a perceptual coordinate system) the path planner can specify it by “pointing” to the world: “Where I look *now* is a new added constraint for the desired path”. Thus the inverse process only needs to be able to designate locomotor goals in oculomotor terms (simpler, known, dynamics than the body-in-vehicle or the vehicle-in-world dynamics).

The forward model (*Localization*) will then be able to compute the path parameters *from current LAF input,* and return the predicted path error for locomotor planning (i.e. generation of the appropriate steering command sequence) to the Path planner, as well as passing it to the lower level as context information for guidance and stabilization. This would be a principled way to understand how (as discussed in Section 3.2) LAFs may be involved in identifying future waypoints which the driver wishes to pass through and setting parameters for the guidance-level control steering commands.

On the forward models side (indicated in **Figure 5C** by top-down and bottom-up arrows), waypoint *Localization*, *Guidance* level target information in online short term memory, and keeping both in check with *Stabilizing* - i.e. gravitoinertial frame and body schema - information could begin to give us an account of how a coherent "image" of visual space is formed (Senders et al., 1967; Land & Furneaux, 1997; Kujala et al., 2016; cf. Loomis et al., 1996; Loomis, Klatzky & Giudice, 2013; cf. Also Aivar et al., 2005; Tatler & Land, 2011; Burr & Morrone, 2012), and in particular provide new ideas on how to think about its contribution to complex real-world behavior in more mechanistic terms. This account raises the possibility of providing the missing links between egocentric orientation and path following and higher-level allocentric navigation – a key topic in computational and cognitive neuroscience (Barry & Burgess, 2014; Spiers & Barry, 2015).

Note that the architecture proposed in **Figure 5C** indeed has coupled inverse and forward models *on all levels*. This is because we find that there is little neurobiological motivation for restricting model-based control to the highest (sequential motor plan) level. Whilst much of this manuscript is focussed on model-based prediction as being primarily useful for "anticipatory" control, there is evidence that forward models may be ubiquitous even at low-levels of the nervous system so it seems reasonable that model-based predictions could have a role across all levels (**Figure 5C**). For example, at the lowest-level forward models use efference copies of motor commands to predict the nature of feedback (reafferent sensory stimulation), explaining the vestibulo-ocular reflex (for reviews see Cullen & Roy, 2004; Cullen, 2004; Angelaki & Hess, 2005) or why one cannot tickle themselves (Blakemore, Wolpert & Frith, 1998). Furthermore, it has been suggested that basic visual tracking, which is known to be able to utilize predictive information about anticipated rather than observed motion (Kowler, 1989; see also Hayhoe et al., 2012), also depends on predictive models (Orban de Xivry et al., 2013; Daye et al., 2014). Finally, saccade planning (relevant to gaze polling saccades between GF and LAF), coordinated eye-hand motor sequences and visual stability are generally assumed to rely on stable transsaccadic memory representations (Aivar et al., 2005; Tatler & Land, 2011; Burr & Morrone, 2012; Loomis, Klatzky & Giudice, 2013).

The fundamental theoretical questions for our multiple waypoint hypothesis concern how locations in the world become waypoints which specify the "desired path". If the process relies on internal models (as proposed above), how rich and enduring is this representation? We are not advocating a metrically accurate euclidean 3D "world model". Our proposal is that the future path is represented by a small number of waypoints (but greater than two) – the "minimal" representation needed to capture the information processing underlying skilled driving and also our interpretation of the "spatial image" or visual buffer" referred to in the studies of short term memory in locomotor control and intermittent visual sampling in driving. Waypoints are identified as oculomotor and locomotor targets by prefrontal motor planning mechanisms, are tracked by posterior parietal updating systems, and associated with long term (e.g. landmark) memory by medial temporal structures and the retrosplenial cortex. The use of these mechanisms in other forms of locomotion is fairly well established in brain research – here we extend their use to driving.

This proposal, if on the right track, calls for a subtle but conceptually deep re-evaluation of the posits of the two-level models that for over 40 years have guided steering control research in psychology, vision science, and engineering. Critically, the two level models as depicted in Figures 1-3 are *travel point* models, not waypoint models (cf. Lappi, 2014), and rest on a near/far distinction (since travel points have fixed distances). With waypoints, however, the near/far distinction becomes less fundamental (a "far" waypoint moves to "near" as one approaches it), and so does the rationale for sticking to just two "levels": if a GF is a fixation that lands on a waypoint in the "far" region, why would not LAFs be fixations that land on waypoints even "further"?

Conceiving of the road preview as multiple waypoints brings an imminent need to analyze the gaze strategies and coordinate transformations needed to establish the egocentric projection of the "desired path". (This is missing in the engineering literature where the desired path and its allocentric to egocentric transformation are assumed to be simply available). On the other hand, it necessitates the incorporation of the short- and long-term memory processes used for maintaining the internal models. (This is missing in the vision science literature where simple control rules are sought that specify actions on the basis of optical variables such as looming, or features of optic flow such as the location of the focus of expansion). We hope we have been able to make a case that this is a promising direction for modelling efforts, with the potential to bring together the engineering, psychological and neuroscience perspectives on driving.

5. Conclusions and Future Directions

We have outlined three steps that we believe will turn out to be instrumental in taking forward visual steering models from the current “two level” framework (**Figure 2B, Figure 3**):

1) treat the coordination of gaze and steering as the core control problem,
2) add (or bring back rather) a third *trajectory planning* level, and
3) assume model-based predictive control at all levels (forward and inverse models coupled via efference copy - but also via the effect steering and eye movements have on sensory feedback)

In **Figure 5C**, we have indicated how these three expansions could be naturally seen to extend the control theoretical framework, allowing vision science to simultaneously incorporate persuasive accounts of intermittency in visual sampling and control, allowing more powerful and complete accounts of driver behaviour. **Figure 5B** outlines in more detail the connections to existing literature in terms of cognitive task analysis of driving and the likely neural basis of those processes (from **Figure 4**).

While we have discussed the problem of visual control of locomotion in the concrete and well-understood case of driving, nothing about the general ideas depends on special aspects of the domain, and therefore have wider relevance to other, less well understood forms of visually guided locomotion. Presented here is a modelling architecture which would support a unified model of driving, by potentially assimilating current state of the art in vision science, engineering and movement neuroscience, and inform future experimental designs.

Whilst “Driving a car” in itself is a behavior of only limited scientific interest (beyond applied science), the task is an excellent example of coordinating multiple fundamental mechanisms (e.g. visual search, object tracking, saccade planning, motor control). One of the unique advantages of the driving domain is that these processes can be examined in highly controlled laboratory experiments or in uncontrolled natural (‘on-road’) environments - and simulators of almost continuously varying degrees of fidelity and complexity in between. This capacity offers vision science the ability to both isolate specific processes and map those processes through to how they are coordinated/interact with one another whilst executing *in situ* behaviours.

Naturalistic tasks – including such as real-world driving or complex virtual reality scenarios – can reveal ecologically valid visual strategies. Their downside is that due to complex stimuli, many-degrees-of-freedom behaviors, and no direct experimental manipulation, interpreting the data in terms of causal dependencies and underlying mechanisms is limited. In restricted laboratory tasks, however, the parameters of stimuli and degrees of freedom allowed in the behavioral response can be controlled. The downside is that imposing such limitations can affect the participants' perceptual and cognitive strategies. Thus, although performance on such (simple) tasks can be more readily rigorously modelled, how well such (simple) models can handle (complex) real-world settings remains unresolved unless correspondence between the strategies in laboratory/simulator and ecological settings can be established.

In developing the conceptual model in Section 4, our purpose is to indicate how research from vision science and engineering research on steering performance, naturalistic eye tracking and brain imaging on driving all do fit together into a coherent picture. Significantly, we made it a strict rule that whenever we invoke some complex process *all* component processes must have been modeled, behaviorally investigated, and their neurobiological underpinnings studied in cognitive or computational neuroscience. So you might say we are waving hands to some extent – but we are not waving them in thin air! Instead, we identify where there are conceptual gaps in the current state-of-the-art modeling paradigms in perceptual psychology and vehicle system dynamics engineering, which could be filled by exploiting cutting-edge work already being done in modeling multisensory perception, eye-hand motor control and navigational wayfinding in neuroscience.

Driving represents one of the few dynamic real-world tasks where a *unified* and *complete* model of task performance may indeed be possible in the near future[9]. Such a unified model of *any* natural task - and driving is a better candidate than most as it is already understood at quite a high level of sophistication and detail - would go a long way to achieving Newell's aim of unified theories of cognition, or, as he put it: "a sufficient theory of a genuine slab of human behavior" (Newell, 1973, p.303). We hope this paper will stimulate engineers, psychologists, neuroscientists and computer scientists to collaborate to develop such a theory.

Acknowledgements

Preparation of the manuscript was made possible by mobility supported by an Erasmus+ agreement between the University of Helsinki and the University of Leeds, the Academy of Finland project Mulsimco (279905, OL), the Experimental Psychology Society (CM), the University of Leeds Research Mobility Programme (CM) and the LS2N École Centrale de Nantes (OL). Discussions with Dr. Richard M. Wilkie, Dr. Gustav Markkula, Dr. Esko Lehtonen and Mr. Jami Pekkanen have been instrumental in developing the ideas. The manuscript was finalized and submitted whilst the autohrs were visiting researchers at École Central de Nantes, hosted by Dr. Franck Mars.

References

Aivar, M. P., Hayhoe, M. M., Chizk, C. L., & Mruczek, R. E. (2005). Spatial memory and saccadic targeting in a natural task. *Journal of Vision*, *5*(3), 177-193.

Allen, R. W., & McRuer, D. (1979). The man/machine control interface—Pursuit control. *Automatica*, *15*(6), 683-686.

Andersen, R. A., Snyder, L. H., Li, C. S., & Stricanne, B. (1993). Coordinate transformations in the representation of spatial information. *Current opinion in neurobiology*, 3(2), 171-176.

Angelaki, D. E., & Hess, B. J. (2005). Self-motion-induced eye movements: effects on visual acuity and navigation. *Nature Reviews Neuroscience, 6*(12), 966-976.

Authié, C.N. & Mestre, D.R.(2012). Path curvature discrimination: dependence on gaze direction and optical flow speed. PLoS ONE 7, e31479.

Ballard, D. H., Hayhoe, M. M., Li, F., Whitehead, S. D., Frisby, J. P., Taylor, J. G., & Fisher, R. B. (1992). Hand-eye coordination during sequential tasks [and discussion]. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *337*(1281), 331-339.

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, *7*(1), 66-80.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*(04), 723-74

Barry, C., & Burgess, N. (2014). Neural mechanisms of self-location. *Current Biology, 24*(8), R330-R339.

Bernardi, G., Cecchetti, L., Handjaras, G., Sani, L., Gaglianese, A., Ceccarelli, R., et al. (2014). It's not all in your car: functional and structural correlates of exceptional driving skills in professional racers. *Frontiers in Human Neuroscience* 8:888.

Billington, J., Field, D. T., Wilkie, R. M., & Wann, J. P. (2010). An fMRI study of parietal cortex involvement in the visual guidance of locomotion. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1495.

Billington, J., Wilkie, R. M., & Wann, J. P. (2013). Obstacle avoidance and smooth trajectory control: neural areas highlighted during improved locomotor performance. *Frontiers in behavioral neuroscience*, *7*, 9.

Blaauw, G. J., Godthelp, H., & Milgram, P. (1984). Optimal control model applications and field measurements with respect to car driving.*Vehicle System Dynamics*, *13*(2), 93-111.

Blakemore, S., Wolpert, D.M. & Frith, C.D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*, 635–640.

Boer, E. R. (1996). Tangent point oriented curve negotiation. In *Intelligent Vehicles Symposium, 1996., Proceedings of the 1996 IEEE* (pp. 7-12).

Boer, E. R. (2016). What Preview Elements do Drivers Need?. *IFAC-PapersOnLine*, *49*(19), 102-107.

Bonner, M. F., & Epstein, R. A. (2017). Coding of navigational affordances in the human visual system. *Proceedings of the National Academy of Sciences*, *114*(18), 4793-4798.

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in cognitive sciences*, *12*, 201–208.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 625-641.

Burr, D. C., & Morrone, M. C. (2012). Constructing Stable Spatial Maps of the Word. *Perception*, *41*(11), 1355-1372.

Calhoun, V. D., Pekar, J. J., McGinty, V. B., Adali, T., Watson, T. D., & Pearlson, G. D. (2002). Different activation dynamics in multiple neural systems during simulated driving. *Human brain mapping*, *16*(3), 158-167.

Cardin, V., & Smith, A. T. (2009). Sensitivity of human visual and vestibular cortical regions to egomotion-compatible visual stimulation. *Cerebral Cortex*, *20*(8), 1964-1973.

Chattington, M., Wilson, M., Ashford, D., & Marple-Horvat, D. E. (2007). Eye–steering coordination in natural driving. *Experimental brain research*, *180*(1), 1-14.

Chatziastros, A., Wallis, G. M., & Bülthoff, H. H. (1999). The effect of field of view and surface texture on driver steering performance. *Vision in Vehicles VII*, 253-259.

Cloete, S. R., & Wallis, G. (2009). Limitations of feedforward control in multiple-phase steering movements. *Experimental brain research*, *195*(3), 481-487.

Cloete, S., & Wallis, G. (2011). Visuomotor control of steering: the artefact of the matter. *Experimental Brain Research*, *208*(4), 475-489.

Cole, D. J., Pick, A. J., & Odhams, A. M. C. (2006). Predictive and linear quadratic methods for potential application to modelling driver steering control. *Vehicle System Dynamics*, *44*(3), 259-284.

Coutton-Jean, C., Mestre, D.R., Goulon, C. & Bootsma, R.J. (2009). The role of edge lines in curve driving. *Transportation Research Part F: Traffic Psychology*, *12*, 483-493.

Crawford, J. D., Henriques, D. Y., & Medendorp, W. P. (2011). Three-dimensional transformations for goal-directed action. *Annual Review of Neuroscience*, *34*, 309-331.

Cullen, K. E. (2004). Sensory signals during active versus passive movement. *Current Opinion in neurobiology*, *14*(6), 698-706.

Cullen, K. E., & Roy, J. E. (2004). Signal processing in the vestibular system during active versus passive head movements. *Journal of Neurophysiology*, *91*(5), 1919-1933.

Daye, P. M., Optican, L. M., Blohm, G., & Lefevre, P. (2014). Hierarchical control of two-dimensional gaze saccades. *Journal of Computational Neuroscience,* 36(3), 355-382.

Desmurget, M., & Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Sciences*, *4*(11), 423-431.

Donges, E. (1978). A two-level model of driver steering behavior. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *20*(6), 691-707.

Dukelow, S. P., DeSouza, J. F., Culham, J. C., van den Berg, A. V., Menon, R. S., & Vilis, T. (2001). Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *Journal of Neurophysiology*, *86*(4), 1991-2000.

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*, 12(10), 388-396.

Epstein, R.A., Vass, L.K. (2014). Neural systems for landmark-based wayfinding in humans. *Philosophical Transactions of the Royal Society B*, *369*: 20120533.

Epstein, R. A., Patai, E. Z., Julian, J. B., & Spiers, H. J. (2017). The cognitive map in humans: spatial navigation and beyond. Nature Neuroscience, *20*(11), 1504-1513.

Franklin, D., & Wolpert, D. (2011). Computational mechanisms of sensorimotor control. Neuron, 72, 425–442.

Fajen, B. R. (2001). Steering toward a goal by equalizing taus. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 953-968.

Fajen, B. R., & Warren, W. H. (2003). Behavioral dynamics of steering, obstable avoidance, and route selection. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 343.

Fajen, B. R. (2013). Guiding locomotion in complex, dynamic environments.*Frontiers in behavioral neuroscience*, *7*, 85, 1-15.

Fernandez-Ruiz, J., Goltz, H. C., DeSouza, J. F., Vilis, T., & Crawford, J. D. (2007). Human parietal "reach region" primarily encodes intrinsic visual direction, not extrinsic movement direction, in a visual–motor dissociation task. *Cerebral Cortex*, *17*(10), 2283-2292.

Field, D. T., Wilkie, R. M., & Wann, J. P. (2007). Neural systems in the visual control of steering. *Journal of Neuroscience*, *27*(30), 8002-8010.

Frissen, I., & Mars, F. (2014). The effect of visual degradation on anticipatory and compensatory steering control. *The Quarterly Journal of Experimental Psychology*, *67*(3), 499-507.

Gawthrop, P., Loram, I., Lakie, M., & Gollee, H. (2011). Intermittent control: a computational theory of human control. *Biological cybernetics*, *104*(1), 31-51.

Gazzaniga, M. S. (2013). Shifting gears: seeking new approaches for mind/brain mechanisms. *Annual review of psychology*, *64*, 1-20.

Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals. *British journal of psychology*, *49*(3), 182-194.

Girard, B., & Berthoz, A. (2005). From brainstem to cortex: computational models of saccade generation circuitry. *Progress in neurobiology, 77*(4), 215-251.

Godthelp, J. (1985). Precognitive control: Open-and closed-loop steering in a lane-change manoeuvre. *Ergonomics*, *28*(10), 1419-1438.

Godthelp, H. (1986). Vehicle control during curve driving. *Human Factors*, *28*(2), 211-221.

Godthelp, J., Blaauw, G. J., & Milgram, P. (1984). Supervisory behaviour in automobile driving: new approaches in modelling vehicle control. SAE Technical Paper 845099, https://doi.org/10.4271/845099

Godthelp, H., Milgram, P., & Blaauw, G. J. (1984). The development of a time-related measure to describe driving strategy. *Human factors*, *26*(3), 257-268.

Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, *3*(1), 49-63.

Hayhoe, M. M., McKinney, T., Chajka, K., & Pelz, J. B. (2012). Predictive eye movements in natural vision. *Experimental Brain Research, 217*(1), 125–136.

Hildreth, E. C., Beusmans, J. M., Boer, E. R., & Royden, C. S. (2000). From vision to action: experiments and models of steering control during driving. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(3), 1106-1132.

Holmes, N. P., & Spence, C. (2004). The body schema and multisensory representation (s) of peripersonal space. *Cognitive processing*, 5(2), 94-105.

Hoult, W. & Cole, D.J. (2008). A neuromuscular model featuring co-activation for use in driver simulation. *Vehicle System Dynamics*, *46*, pp. 175-189.

Huang, R. S., Chen, C. F., & Sereno, M. I. (2015). Neural substrates underlying the passive observation and active control of translational egomotion. *Journal of Neuroscience*, *35*(10), 4258-4267.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience*, *22*(16), 7195-7205.

Itkonen, T., Pekkanen, J., & Lappi, O. (2015). Driver gaze behavior is different in normal curve driving and when looking at the tangent point. *PloS one*, *10*(8), e0135505.

Jeon, H. A. (2014). Hierarchical processing in the prefrontal cortex in a variety of cognitive domains. *Frontiers in systems neuroscience*, *8* (223): 1-8.

Johns, T. A., & Cole, D. J. (2015). Measurement and mathematical model of a driver's intermittent compensatory steering control. *Vehicle System Dynamics*, *53*(12), 1811-1829.

Johnson, L., Sullivan, B., Hayhoe, M., & Ballard, D. (2014). Predicting human visuomotor behaviour in a driving task. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369*(1636), 20130044.

Kan, K., Schweizer, T. A., Tam, F., and Graham, S. J. (2013). Methodology for functional MRI of simulated driving. *Medical Physics 40*, 012301.

Keele, S. W. (1968). Movement control in skilled motor performance.*Psychological Bulletin*, *70*(6 Part 1), 387-403.

Keen, S. D., & Cole, D. J. (2006). Steering control using model predictive control and multiple internal models. In: Che-Wun Hong and Guojia-Kexue-Weiyuanhui (editors): *Proceedings of the 8th International Symposium on Automotive Control* (pp. 599–604). Hsinchu National Tsing Hua University. Taipei, Taiwan.

Keen, S. D., & Cole, D. J. (2011). Application of time-variant predictive control to modelling driver steering skill. *Vehicle System Dynamics*, *49*(4), 527-559.

Keen, S. D., & Cole, D. J. (2012). Bias-free identification of a linear model-predictive steering controller from measured driver steering behavior. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *42*(2), 434-443.

Kim, N. G., & Turvey, M. T. (1999). Eye movements and a rule for perceiving direction of heading. *Ecological Psychology*, *11*(3), 233-248.

Kleinman, D., Baron, S., & Levison, W. (1971). A control theoretic approach to manned-vehicle systems analysis. *IEEE Transactions on Automatic Control*, *16*(6), 824-832.

Koechlin, E., Ody, C., and Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science,* 302, 1181–1185.

Koenderink, J. J. (1986). Optic flow. *Vision research*, *26*(1), 161-179.

Kountouriotis, G. K., Mole, C. D., Merat, N., & Wilkie, R. M. (2016). The need for speed: global optic flow speed influences steering. *Royal Society open science*, *3*(5), 160096.

Kowler, E. (1989). Cognitive expectations, not habits, control anticipatory smooth oculomotor pursuit. *Vision research*, *29*(9), 1049-1057.

Kowler, E. (2011). Eye movements: The past 25years. *Vision research*, *51*(13), 1457-1483.

Krauzlis, R. J. (2004). Recasting the smooth pursuit eye movement system. *Journal of neurophysiology*, *91*(2), 591-603.

Kujala, T., Mäkelä, J., Kotilainen, I., & Tokkonen, T. (2016). The Attentional Demand of Automobile Driving Revisited Occlusion Distance as a Function of Task-Relevant Event Density in Realistic Driving Scenarios. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *58*(1), 163-180.

Land, M.F. (1992). Predictable eye-head coordination during driving. *Nature*, *359*(6393), 318-320.

Land, M.F. (1998). The visual control of steering. In: Harris, L.R. & Jenkin, M. (eds.): *Vision and action*. Cambridge university press: Cambridge. Pp. 163-180.

Land, M. F. (2007). Fixation strategies during active behavior: A brief history. In: R.P.G: vanGompel, M.H.Fischer, W.S. Murray & R.L.Hill (eds.): *Eye movements: A window on mind and brain*, 75-95. Amsterdam: Elsevier.

Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *352*(1358), 1231-1239.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities?. *Vision research*, *41*(25), 3559-3565.

Land, M., & Horwood, J. (1995). Which parts of the road guide steering? *Nature*, *377*(6547), 339.

Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature*, *369*(6483), 742-744.

Land, M. F., Mennie, N., & Rusted, J. (1999). Eye movements and the roles of vision in activities of daily living: making a cup of tea. *Perception*, *28*(4), 1311-1328.

Land, M. F., & Tatler, B. W. (2001). Steering with the head: The visual strategy of a racing driver. *Current Biology*, *11*(15), 1215-1220.

Lappi, O. (2014). Future path and tangent point models in the visual control of locomotion in curve driving. *Journal of vision*, *14*(12), 21-21.

Lappi, O. (2015). The Racer's Brain–How Domain Expertise is Reflected in the Neural Substrates of Driving. *Frontiers in human neuroscience*, *9*, 635

Lappi, O. (2016). Eye movements in the wild: oculomotor control, gaze behavior & frames of reference. *Neuroscience & Biobehavioral Reviews*, *69*, 49-68.

Lappi, O., & Lehtonen, E. (2013). Eye-movements in real curve driving: pursuit-like optokinesis in vehicle frame of reference, stability in an allocentric reference coordinate system. *Journal of Eye Movement Research*, *6*(1), 1-13.

Lappi, O., Lehtonen, E., Pekkanen, J., & Itkonen, T. (2013). Beyond the tangent point: gaze targets in naturalistic driving. *Journal of vision*, *13*(13):11, 1-18.

Lappi, O., Pekkanen, J., & Itkonen, T. H. (2013). Pursuit eye-movements in curve driving differentiate between future path and tangent point models. *PloS one*, *8*(7), e68326.

Lappi, O., Rinkkala, P. & Pekkanen, J. (2017). Systematic Observation of An Expert Driver's Gaze Strategy - An On-Road Case Study. *Frontiers in Psychology*. *8*:620.

Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, *5*(4), 437-459.

Lee, D. N., Bootsma, R. J., Land, M., Regan, D., & Gray, R. (2009). Lee's 1976 paper. *Perception*, *38*(6), 837-858.

Lee, D. N., & Kalmus, H. (1980). The optic flow field: The foundation of vision [and discussion]. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *290*(1038), 169-179.

Lehtonen, E., Lappi, O., Kotkanen, H., & Summala, H. (2013). Look-ahead fixations in curve driving. *Ergonomics*, *56*(1), 34-44.

Lehtonen, E., Lappi, O., Koirikivi, I., & Summala, H. (2014). Effect of driving experience on anticipatory look-ahead fixations in real curve driving. *Accident Analysis & Prevention*, *70*, 195-208.

Li, L., & Cheng, J. C. K. (2011). Heading but not path or the tau-equalization strategy is used in the visual control of steering toward a goal. *Journal of Vision, 11*(12), 1-12.

Li, L., & Cheng, J. C. K. (2013). Visual strategies for the control of steering toward a goal. *Displays*, *34*(2), 97-104.

Li, Y. O., Eichele, T., Calhoun, V. D., and Adali, T. (2012). Group study of simulated driving fMRI data by multiset canonical correlation analysis. *Journal of Signal Processing Systems. 68*, 31–48.

Loomis, J. M., Da Silva, J. A., Philbeck, J. W., & Fukusima, S. S. (1996). Visual perception of location and distance. *Current Directions in Psychological Science*, *5*(3), 72-77.

Loomis, J. M., Klatzky, R. L., & Giudice, N. A. (2013). Representing 3D space in working memory: Spatial images from vision, hearing, touch, and language. In Lacey, S. & Lawson, R. (editors): *Multisensory Imagery* (pp. 131-155). Springer New York.

MacAdam, C. (2001). *Development of a driver model for near/at-limit vehicle handling* (No. UMTRI-2001-43,).

Macadam, C. C. (2003). Understanding and modeling the human driver. *Vehicle System Dynamics*, *40*(1-3), 101-134.

Macuga, K. L., Beall, A. C., Kelly, J. W., Smith, R. S., & Loomis, J. M. (2007). Changing lanes: Inertial cues and explicit path information facilitate steering performance when visual feedback is removed. *Experimental brain research*, *178*(2), 141-150.

Markkula, G., Boer, E., Romano, R., & Merat, N. (2018). Sustained sensorimotor control as intermittent decisions about prediction errors: Computational framework and application to ground vehicle steering. *Biological Cybernetics* (in press). https://doi.org/10.1007/s00422-017-0743-9

Markkula, G. (2014). Modeling driver control behavior in both routine and near-accident driving. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 58, No. 1, pp. 879-883). Sage CA: Los Angeles, CA: SAGE Publications.

Marple-Horvat, D., Chattington, M., Anglesea, M., Ashford, D.G., Wilson, M. & Keil, D. (2005). Prevention of coordinated eye movements and steering impairs driving performance. *Experimental Brain Research*, *163*, 411–420.

Mars, F. (2008). Driving around bends with manipulated eye-steering coordination. *Journal of vision*, *8*(11):10, 1-11.

Mars, F. & Chevrel, P. (2017). Modelling human control of steering for the design of advanced driver assistance systems, *Annual Reviews in Control*, *44*, p. 292-302.

Mars, F., & Navarro, J. (2012). Where we look when we drive with or without active steering wheel control. *PloS one*, *7*(8), e43858.

Mars, F., Saleh, L., Chevrel, P., Claveau, F. & Lafay J.-F. (2011). Modeling the visual and motor control of steering with an eye to shared-control automation. *Proceedings of the human factors and ergonomics society annual meeting*, *55,* pp. 1422-1426.

McGuire, L.M., Sabes, P.N., 2009. Sensory transformations and the use of multiplereference frames for reach planning. *Nature Neuroscience*, *12*(8), 1056–1061.

McRuer, D. T., Allen, R. W., Weir, D. H., & Klein, R. H. (1977). New results in driver steering control models. *Human factors*, *19*(4), 381-397.

McRuer, D. T., Graham, D., Krendel, E. S., & Reisener Jr, W. (1965). Human pilot dynamics in compensatory systems. theory, models and experiments with controlled element and forcing function variations. *Wright-Patterson AFB (OH): Air Force Flight Dynamics Laboratory*.

McRuer, D. T., & Jex, H. R. (1967). A review of quasi-linear pilot models. *IEEE Transactions on Human Factors in Electronics*, (3), 231-249.

McRuer, D.T. & Krendel, e.s. (1957). Dynamic Response of Human Operators. WADC Technical report 56-524. Wright Air Development Center, Ohio.

McRuer, D.T. & Krendel, E.S. (1959). The human operator as a servo system element, *Journal of the Franklin Institute*, *267*(5), 381-403.

McRuer, D. T., Weir, D. H., Jex, H. R., Magdaleno, R. E., & Allen, R. W. (1975). Measurement of driver-vehicle multiloop response properties with a single disturbance input. *IEEE Transactions on Systems, Man, and Cybernetics*, (5), 490-497.

McRuer, D. (1980). Human dynamics in man-machine systems. *Automatica*,*16*(3), 237-253.

Meister, M. L., & Buffalo, E. A. (2016). Getting directions from the hippocampus: the neural connection between looking and memory. *Neurobiology of learning and memory*, 134, 135-144.

Mennie, N., Hayhoe, M., & Sullivan, B. (2007). Look-ahead fixations: anticipatory eye movements in natural tasks. *Experimental Brain Research*,*179*(3), 427-442.

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural networks*, *9*(8), 1265-1279.

Miall, R. C., Reckess, G. Z., & Imamizu, H. (2001). The cerebellum coordinates eye and hand tracking movements. *Nature neuroscience, 4*(6), 638-644.

Mitchell, A., Czajkowksi, R., Zhang, N., Jeffery, K. & Nelson, A. (2017). Retrosplenial cortex and its role in spatial cognition. bioRxiv preprint Sep. 19, 2017; doi: http://dx.doi.org/10.1101/190801.

Mole, C. D., Kountouriotis, G., Billington, J., & Wilkie, R. M. (2016). Optic flow speed modulates guidance level control: New insights into two-level steering. *Journal of experimental psychology: human perception and performance, 42(*11), 1818-1838.

Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., & Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nature neuroscience*, 3(12), 1322-1328.

Munoz, D. P. (2002). Commentary: saccadic eye movements: overview of neural circuitry. *Progress in brain research*, 140, 89-96.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–308). New York: Academic Press.

Nummenmaa, L., Oksama, L., Glerean, E., & Hyönä, J. (2017). Cortical Circuit for Binding Object Identity and Location During Multiple-Object Tracking. *Cerebral Cortex*, *27*(1), 162-172.

Näätänen, R., & Summala, H. (1974). A model for the role of motivational factors in drivers' decision-making. *Accident Analysis & Prevention*, *6*(3), 243-261.

Näätänen, R., & Summala, H. (1976). *Road-user Behaviour and Traffic Accidents*. Amsterdam*;* North-Holland Publishing Company.

Nash, C. J., Cole, D. J., & Bigler, R. S. (2016). A review of human sensory dynamics for application to models of driver steering and speed control.*Biological cybernetics*, *110*(2-3), 91-116.

Odhams, A. M. C., & Cole, D. J. (2009). Application of linear preview control to modelling human steering control. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, *223*(7), 835-853.

Odhams, A. M. C., & Cole, D. J. (2010). Identification of a driver's preview steering control behaviour using data from a driving simulator and a randomly curved road path. In *Proc. 10th Int. Symp. AVEC* (pp. 605-609).

Orban de Xivry, J. J., Coppe, S., Blohm, G., & Lefevre, P. (2013). Kalman filtering naturally accounts for visually guided and predictive smooth pursuit dynamics. *The Journal of Neuroscience*, 33(44), 17301-17313.

Park, S., & Chun, M. M. (2009). Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *Neuroimage*, 47(4), 1747-1756.

Pekkanen, J., Lappi, O., Itkonen, T. H., & Summala, H. (2017). Task-difficulty homeostasis in car following models: experimental validation using self-paced visual occlusion. *PLoS one*, *12*(1), e0169704.

Pelz, J. B., & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision research*, *41*(25), 3587-3596.

Peuskens, H., Sunaert, S., Dupont, P., Van Hecke, P., & Orban, G. A. (2001). Human brain regions involved in heading estimation. *Journal of Neuroscience, 21*(7), 2451-2461.

Pierrot-Deseilligny, C., Milea, D., & Müri, R. M. (2004). Eye movement control by the cerebral cortex. *Current opinion in neurology*, *17*(1), 17-25.

Pierrot-Deseilligny, C., Müri, R. M., Ploner, C. J., Gaymard, B., & Rivaud-Pechoux, S. (2003). Cortical control of ocular saccades in humans: a model for motricity. *Progress in brain research, 142*, 3-17.

Pitzalis, S., Sereno, M. I., Committeri, G., Fattori, P., Galati, G., Patria, F., & Galletti, C. (2009). Human V6: the medial motion area. *Cerebral Cortex*, *20*(2), 411-424.

Pitzalis, S., Sdoia, S., Bultrini, A., Committeri, G., Di Russo, F., Fattori, P., ... & Galati, G. (2013). Selectivity to translational egomotion in human brain motion areas. *PLoS ONE*, *8*(4), e60241.

Pouget, A., Deneve, S., & Duhamel, J. R. (2002). A computational perspective on the neural basis of multisensory spatial representations. Nature Reviews Neuroscience, *3*(9), 741-747.

Prokop, G. (2001). Modeling human vehicle driving by model predictive online optimization. *Vehicle System Dynamics*, *35*(1), 19–53.

Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nature reviews neuroscience, 5*(3), 184-194.

Regan, D., & Beverley, K. I. (1982). How do we avoid confounding the direction we are looking and the direction we are moving? *Science*, 215(4529), 194-196.

Regan, D., & Gray, R. (2000). Visually guided collision avoidance and collision achievement. *Trends in Cognitive Sciences*, *4*(3), 99-107.

Saleh, L., Chevrel, P., Mars, F., Lafay, J. F., & Claveau, F. (2011). Human-like cybernetic driver model for lane keeping. *IFAC Proceedings*, *44*(1), 4368-4373.

Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture.*Human Factors: The Journal of the Human Factors and Ergonomics Society*, *48*(2), 362-380.

Salvucci, D. D., & Gray, R. (2004). A two-point visual control model of steering. *Perception*, *33*(10), 1233-1248.

Salvucci, D. D., Zuber, M., Beregovaia, E., & Markley, D. (2005). Distract-R: Rapid prototyping and evaluation of in-vehicle interfaces. In *Human Factors in Computing Systems: CHI 2005 Conference Proceedings* (pp. 581–589). New York, NY: ACM Press.

Schepers, J. P., Den Brinker, B. P. L. M., De Waard, D., Twisk, D. A. M., Schwab, A. L., & Smeets, J. B. (2013). Studying the role of vision in cycling: Critique on restricting research to fixation behaviour. *Accident Analysis & Prevention*, *59*, 466-468.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning.*Psychological review*, *82*(4), 225.

Schmidt, R. A. (2003). Motor schema theory after 27 years: reflections and implications for a new theory. *Research quarterly for exercise and sport*, *74*(4), 366-375.

Schweizer, T. A., Kan, K., Hung, Y., Tam, F., Naglie, G., & Graham, S. (2013). Brain activity during driving with distraction: an immersive fMRI study. Frontiers in human neuroscience, 7, 53, 1-11.

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).

Sentouh, C., Chevrel, P., Mars, F., & Claveau, F. (2009, October). A sensorimotor driver model for steering control. In *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on* (pp. 2462-2467). IEEE.

Sharp, R. S. (2005). Driver steering control and a new perspective on car handling qualities. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, *219*(10), 1041-1051.

Sharp, R.S., Casanova, D. & Symonds, P. (2000). A Mathematical Model for Driver Steering Control, with Design, Tuning and Performance Results, *Vehicle System Dynamics*, *33*(5), 289-326.

Sharp, R. S., & Peng, H. (2011). Vehicle dynamics applications of optimal control theory. *Vehicle System Dynamics*, *49*(7), 1073-1111.

Sentouh, C., Chevrel, P., Mars, F., & Claveau, F. (2009). A sensorimotor driver model for steering control. *IEEE International Conference on Systems, Man and Cybernetics, SMC 2009,* 2462-2467.

Sharp, R. S., & Valtetsiotis, V. (2001). Optimal preview car steering control.*Vehicle System Dynamics*, *35*(SUPPL. 1), 101-117.

Shea, C. H., & Wulf, G. (2005). Schema theory: A critical appraisal and reevaluation. *Journal of Motor Behavior*, *37*(2), 85-102.

Shinar, D., McDowell, E. D., & Rockwell, T. H. (1977). Eye movements in curve negotiation. *Human Factors*, *19*(1), 63-71.

Smith, A. T., Wall, M. B., Williams, A. L., & Singh, K. D. (2006). Sensitivity to optic flow in human cortical areas MT and MST. *European Journal of Neuroscience*, *23*(2), 561-569.

Smith, A. T., Wall, M. B., & Thilo, K. V. (2011). Vestibular inputs to human motion-sensitive visual cortex. *Cerebral cortex*, *22*(5), 1068-1077.

Spiers, H. J., & Barry, C. (2015). Neural systems supporting navigation. *Current Opinion in Behavioral Sciences*, *1*, 47-55.

Spiers, H. J., and Maguire, E. A. (2006). Thoughts, behaviour, and brain dynamics during navigation in the real world. *Neuroimage, 31*, 1826–1840.

Spiers, H. J., and Maguire, E. A. (2007a). Neural substrates of driving behaviour. *Neuroimage, 36*, 245–255.

Spiers, H. J., & Maguire, E. A. (2007b). The neuroscience of remote spatial memory: a tale of two cities. *Neuroscience*, *149*(1), 7-27.

Summers, J. J., & Anson, J. G. (2009). Current status of the motor program: Revisited. *Human Movement Science*, *28*(5), 566-577.

Tanner, J., & Itti, L. (2017). Goal relevance as a quantitative model of human task relevance. *Psychological review*, *124*(2), 168-178.

Tatler, B. W., & Land, M. F. (2011). Vision and the representation of the surroundings in spatial memory. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *366*(1564), 596-610.

Timings, J. P., & Cole, D. J. (2012). Vehicle trajectory linearisation to enable efficient optimisation of the constant speed racing line. *Vehicle System Dynamics*, *50*(6), 883-901.

Underwood, G., Chapman, P. et al. (1999). The visual control of steering and driving: Where do we look when negotiating curves? In: A.G.Gale, I.D.Brown, C.M.Haslegrave and S.P.Taylor (eds.) V*ision in Vehicles VII.* Amsterdam, Elsevier.

Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience*, *10*(11), 792-802.

Vansteenkiste, P., Cardon, G., D'Hondt, E., Philippaerts, R., & Lenoir, M. (2013). The visual control of bicycle steering: The effects of speed and path width. *Accident Analysis & Prevention*, *51*, 222-227.

Wall, M. B., & Smith, A. T. (2008). The representation of egomotion in the human brain. *Current biology*, 18(3), 191-194.

Wallis, G., Chatziastros, A., & Bülthoff, H. (2002). An unexpected role for visual feedback in vehicle steering control. *Current Biology*, *12*(4), 295-299.

Wallis, G., Chatziastros, A., Tresilian, J., & Tomasevic, N. (2007). The role of visual and nonvisual feedback in a vehicle steering task. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1127.

Walter, H., Vetter, S. C., Grothe, J. O., Wunderlich, A. P., Hahn, S., & Spitzer, M. (2001). The neural correlates of driving. *Neuroreport*, *12*(8), 1763-1767.

Wann, J., & Land, M. (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., & 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. 401-419). Kluwer: Dordrecht.

Warren, W. H. (2006). The dynamics of perception and action. *Psychological review*, *113*(2), 358.

Warren, W. H. (2007). Action-scaled information for the visual control of locomotion. In Pepping, G-j and Grealy, M.A. (editors): *Closing the gap: The scientific writings of David N. Lee*, pp. 243-258.

Wierwille, W. W., Gagne, G. A., & Knight, J. R. (1967). An experimental study of human operator models and closed-loop analysis methods for high-speed automobile driving. *IEEE Transactions on Human Factors in Electronics*, (3), 187-201.

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*(23), 2014-2017.

Wilkie, R. M., & Wann, J. P. (2003). Eye-movements aid the control of locomotion. *Journal of Vision*, *3*(11), 3-3.

Wilkie, R. M., & Wann, J. P. (2006). Judgments of path, not heading, guide locomotion. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(1), 88-96.

Wilkie, R. M., Wann, J. P., & Allison, R. S. (2008). Active gaze, visual look-ahead, and locomotor control. *Journal of experimental psychology: Human perception and performance*, *34*(5), 1150-1164.

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is.

Wilson, M., Stephenson, S., Chattington, M., & Marple-Horvat, D. E. (2007). Eye movements coordinated with steering benefit performance even when vision is denied. *Experimental brain research*, *176*(3), 397-412.

Wilson, M., Chattington, M., & Marple-Horvat, D. E. (2008). Eye movements drive steering: Reduced eye movement distribution impairs steering and driving performance. *Journal of motor behavior*, *40*(3), 190-202.

Wolbers, T., Hegarty, M., Büchel, C., & Loomis, J. M. (2008). Spatial updating: how the brain keeps track of changing object locations during observer motion. *Nature neuroscience*, *11*(10), 1223-1230.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it?. *Nature reviews neuroscience*, *5*(6), 495-501.

Wolpert, Daniel M., Jörn Diedrichsen, and J. Randall Flanagan. "Principles of sensorimotor learning." *Nature Reviews Neuroscience* 12, no. 12 (2011): 739-751.

Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural networks*, *11*(7), 1317-1329.

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *nature neuroscience*, *3*, 1212-1217.

Zhao, H., & Warren, W. H. (2015). On-line and model-based approaches to the visual control of action. *Vision research*, *110*, 190-202.

Footnotes

1 WHO statistics give a number of 1.78 *billion* vehicles registered for road use at the end of 2015 (http://apps.who.int/gho/data/node.main.A995; retrieved 2017/03/04). This does not include vehicles in military or sports use.

2 There was discussion of the three-level McRuer et al. steering control framework in traffic psychology still in the 1980s, with regard to intermittency in lane keeping, lane change and curve driving tasks (Blaauw, Godthelp & Milgram, 1984; Godthelp, Milgram & Blaauw, 1984; Godthelp, 1985, 1986; see section 3 Internal models and predictive control in driving), but more recently traffic psychology seems to develop ideas in the "two-level framework" (e.g. Vansteenkiste et al., 2013; Schepers et al.. 2013).

3 These are optical variables which - under some idealizing assumptions - *specify* one's locomotor state or the error between current state and a desired "target" state in terms of visual angles. Responding to optical variables *directly* means that no inferential process constructing an internal representation of the 3D layout would be involved - instead "active vision" is a problem of *coordinating* movement with changes in specific optical variables, i.e. coupling of vision and action. This is usually contrasted with *integrating* the information from multiple variables into a "cognitive" representation.

4 Although the replicability of some of Land & Horwood's (1995) observations (such as a half-way "optimum") has been challenged (Chatziastros, Wallis & Bülthoff, 1999; Cloete & Wallis, 2011), the consensual view is that the basic idea of the differential effects of the visual availability of information from near vs. far regions the road view is sound (cf. Frissen & Mars, 2014; Mole et al., 2016).

5 Some models (e.g. Wilkie et al., 2008) have response damping and inertia in the neuromuscular system designed into the model architecture, providing the model with a capacity to deal with intermittent/coarse input (Wann & Wilkie, 2004) without going significantly off-track. However, it is still an open question whether and to what extent feedback steering control can robustly cope with intermittent or unreliable input with more complex road geometry.

6 Wallis et al. did no eye tracking so only temporal limits can be discussed on basis of their data, but a natural interpretation would be that the prediction horizon extend to the current point of fixation, i.e. about 15m at 50 km/h (or 1.5m at walking speed), assuming 2s TH typical of GF.

7 Note that in naturalistic eye tracking, interpreting what the intended gaze target is - especially when the plain road surface often has no distinctive objects or visual features and the putative steering points can fall in close optical proximity - is fraught with difficulty (see Lappi 2014 for a methodological critique of early guiding fixation studies overinterpreting gaze in the vicinity of the theoretically preferred target as evidence). For the purposes of this section we only need to assume that look-ahead fixations fall on the road surface and substantially further than the typical guiding fixation - an uncontroversial assumption.

8 Specifically, Pelz & Canosa (2001) observed that as the subject approached the sink, in addition to the tap (relevant for the next action) a fixation was also made to the soap dispenser (relevant for a

future action), then gaze returned to the tap while washing, until switching again to the soap dispenser before reaching for it.

9 Complete in the sense that the model should i. perform at or near human level in a *variety* of scenarios that is representative of the variety of tasks contexts people actually face in the real world, ii. reproduce not only individual performance measures (such as reaction times or steering error) but predict other measures indicative of the strategy humans actually employ in the wild (e.g. eye movement patterns, brain activity), iii. capture performance and strategies at different levels of skill from novice to expert.