Future path and tangent point models in the visual control of locomotion in curve driving

Otto Lappi

Cognitive Science Division/Traffic Research Unit, Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland

Introduction

Common sense suggests that drivers should “keep their eyes on the road” and “look where they are going.” This seems especially good advice when steering in a bend. Indeed, measurement of drivers’ eye movements has shown that during curve negotiation gaze is focused in a remarkably small region in the visual field, anticipating the changing direction of locomotion. But where exactly in the road scene does gaze fall during curve driving? And how do eye movements contribute to steering a path through the curve?

For the past 20 years, the point of departure for on-road studies of visual behavior in curve driving has been the seminal paper published by Land and Lee (1994). In that paper, a model of a visual steering strategy was introduced, based on the use of the tangent point (TP). This is a point on the lane edge on the inside of the curve where the line of sight is tangential to the lane edge, and corresponds to the point in the driver’s visual field where the visual orientation of the projection of the edgeline is reversed (Figure 1A). TP strategies (tracking the TP in the visual field and using the TP visual direction to control steering on-line) quickly became the default account of how vision is used during curve negotiation. More recent studies, however, have questioned the generality of the TP hypothesis and presented theoretical and empirical arguments to the effect that instead of or in addition to looking at the TP, drivers seek out target points on the road surface that they desire their locomotor trajectory to fall on—their future path (FP).

This paper reviews theoretical and empirical work done in the past 20 years on these TP and FP models of the visual control of locomotion in curve driving. The focus is on field studies that have explicitly assessed the TP or the FP hypothesis. In order to keep the review to a reasonable length while at the same time covering the topic in some depth, the choice has been made to limit the scope to on-road studies of visual behavior in curve driving. Perhaps the most comprehensive overview of the variety of different visual behaviors in driving to date is Land and Tatler (2009, chapter 7; see also Land, 1998, 2006, 2007). The present paper focuses on the specific subtask of visual control of steering in bends, covering it in more depth. For the specialist, a synopsis of empirical results from on-road TP and FP studies are given in a separate appendix—in the main text the discussion focuses on how the data to date relate to the different theories.
Figure 1. (A) TPs and FP reference points in the visual field. The visual field is partitioned into a near zone and a far zone at the level of the TP at the lane edge (Land, 1998; Salvucci & Gray, 2004), the point where the visual angle of the lane edge reverses. Note that there are multiple TPs, and the TP on the lane edge is chosen by convention. Lane edges are visually demarcated by painted edge lines and a centerline (if present). NP_{FP} and NP_{LE} refer to near points on the FP and on the lane edge, respectively (Salvucci & Gray, 2004). FP_A = FP travel point adjacent to the TP (Boer, 1996), FP_B = FP travel point beyond the TP (Wann & Land, 2000), and TP_{LE} = tangent
point at the lane edge (Land & Lee, 1994). Road edge refers to the physical edge of the paved surface. The curve apex refers to the physical region on the inside of the bend. (B) Schematic illustration of the complex flow created by a curvilinear trajectory, with waypoints (WP_a, WP_b) adjacent to and beyond the TP. Note that a waypoint fixed in the three-dimensional scene will follow the local visual flow, whereas the TP, FP1, and FP2 are travel points that remain stable in the visual field but move in the three-dimensional scene with the observer. All TPs (including the TP on the lane edge; Land & Lee, 1994) fall on a zero horizontal flow equal flow cylinder: The local flow at every TP is vertical (Raviv & Herman, 1991; see Figure 3).

Although some of the issues are a bit technical, the exposition aims to be straightforward and to make the issues accessible and interesting to the nonspecialist and the arguments as intuitive as possible without oversimplifying the complexity of driver visual behavior or the models invoked to explain it.

The paper is organized as follows: “Driving as real-world visual behavior” outlines the context from which the visual control of locomotion is here approached—that is, research on visual strategies in naturalistic tasks. This line of research has been progressively developing over the past 25 years and has been recently reviewed by, for example, Regan and Gray (1999), Hayhoe and Ballard (2005), Land (2006, 2007), Kowler (2011), Tatler and Land (2011), and Tatler, Hayhoe, Land, and Ballard (2011). “Visual strategies in curve driving” outlines TP or FP models and their predictions for gaze behavior. The choice has been made to arrange the models in terms of what they postulate as the gaze target—TP versus points on the FP—and whether they describe the visual stimulus in terms of monitoring the egocentric visual direction (VD) of steering points or the direction of local visual flow relative to allocentric vertical (however, as we shall see, many of the issues are less a clear-cut than simple dichotomous alternatives). “On-road studies of gaze behavior during curve negotiation” covers empirical on-road studies of real driving spanning the 20-year interval (1994–2014) from the first on-road TP study to the present time. “Visual strategies in curve driving” and “On-road studies of gaze behavior during curve negotiation” are not a comprehensive historical overview of driver models. For example, longitudinal control and lane changing models are not discussed, and only field studies are covered systematically. Laboratory and simulator studies are discussed where they are directly relevant to the interpretation of on-road data. While on-road studies can claim ecological validity, it is often at the cost of experimental control and formidable methodological challenges that make the data difficult to interpret in terms of underlying visual mechanisms. The complementary strengths and weaknesses of laboratory and field research are discussed in “Conclusions: The next 20 years.”

Driving as real-world visual behavior

For the visual scientist, the major scientific interest in driving lies in understanding the organization of the visuomotor strategies involved and what they can reveal about the organization of visuomotor skills generally. An understanding of the neural basis of the visual, attentional, and motor systems responsible for controlling gaze, the body—and in vehicle-assisted locomotion, the vehicle—is beginning to emerge from studies of eye movements in everyday tasks in naturalistic conditions (e.g., Land, 1992; Land, Mennie, & Rusted, 1999; Hayhoe, Shrivastava, Mruczek, & Pelz, 2003; Mennie, Hayhoe, & Sullivan, 2007; Foulsham, Walker, & Kingstone, 2011) as well as recent brain imaging studies (Walter et al., 2001; Jeong et al., 2006; Billington, Field, Wilkie, & Wann, 2010; Field, Wilkie, & Wann, 2007). Still, it remains to be elucidated exactly what the relevant brain systems are, how they represent visual space, and what computations coordinate the representation(s) of space with the attentional and oculomotor systems responsible for overt visual attention (gaze allocation) and the representation(s) of the movements of the body and vehicle (Tatler & Land, 2011). But while the details of how the brain represents visuospatial, motor, and temporal information remain unknown, research on eye movements in naturalistic tasks has nevertheless uncovered some general qualitative principles of overt gaze behavior in active motor control “in the wild” (Regan & Gray, 1999; Hayhoe & Ballard, 2005; Land, 2006, 2007; Kowler, 2011; Tatler & Land, 2011; Tatler et al., 2011). Each can be readily applied to the domain of driving.

1. **Gaze behavior is stereotypical (repeatable within the same task and context) both within and across individuals and across studies.** In curve driving, this is exemplified by the visual orientation to the curve apex region (“TP orientation”; Land & Lee, 1994) and the pattern of optokinetic nystagmus elicited in curve driving (Authié & Mestre, 2012; Lappi & Lehtonen, 2013; Lappi, Pekkanen, & Itkonen, 2013b).

2. **Most fixations are task governed (“top down” rather than driven by visual saliency).** Fixations are typically focused on immediately task-relevant targets and
locations, monitoring and guiding ongoing action, and checking the outcomes of behavior. Again, TP orientation or FP orientation (generally, “looking where you are going”) accounts for a large percentage of gaze time in natural locomotion and, in particular, in curves, where steering needs to be precisely matched to curve geometry.

3. Individual fixations have identifiable functional roles that can be explained at the level of task analysis, though the explanations are sometimes surprising or unintuitive. In the curve negotiation literature, this principle is well exemplified by the retinal flow theories of Kim and Turvey (1999) and Wann and Swapp (2000), reviewed below. Here, eye movements and steering are jointly controlled to produce linear retinal flow lines. This pattern, and the way it conveys information about the FP and steering error, can be understood only from detailed geometrical analyses at the level of individual fixations. These patterns are not immediately accessible to self-observation by experienced drivers.

4. Eye movements are temporally coupled to the information requirement of the specific task phase. In these so-called guiding fixations (Mennie, Hayhoe, & Sullivan, 2007), gaze leads action by about 1 to 2 s and disengages before completion (gaze switch to a new target) unless the task requires continuous monitoring or tracking. In driving, the gaze-to-steering lag is typically about 1 to 2 s (Land, 1992; Land & Tatler, 2001; Chattington, Wilson, Ashford, & Marple-Horvat, 2007), and the time headway (from current location to the point of fixation) is typically about 2 s (Lappi & Lehtonen, 2013; Lehtonen, Lappi, Koirikivi, & Summala, 2014).

5. When possible, use of short-term memory (encoding and retrieval) tends to be minimized. If possible within task and time limitations—and in terms of sufficiently accurate performance—information is picked up as needed, “just in time” (Ballard, Hayhoe, & Pelz, 1995). This may reflect the inherent limitations of memory encoding and retrieval speed and reliability and storage capacity (properties of the human cognitive architecture) or an efficient strategy of using “cheap” eye movements to pick up information that would be more effortful to encode and retrieve from memory. With experience, many skills become more automatized, reducing the dependence on slow and laborious “controlled processing” dependent on working memory. For example, with experience drivers can begin to share gaze time and memory resources between the locomotor task and other, secondary tasks such as adjusting the radio, indicating that the basic tasks of steering and speed selection are now running with a minimal working memory load.

6. Preview behavior in skilled behavior anticipates upcoming subtasks by interleaving look-ahead fixations to objects and locations that will be relevant in the future in between the more prevalent just-in-time fixations or guiding fixations. While very long preview times (several seconds) are an exception, not the norm, anticipatory glances to objects or locations relevant for upcoming task phases do occasionally occur. In driving, this is seen in look-ahead fixations or gaze polling (Wilkie, Wann, & Allison, 2008) away from the typical guiding fixation region (in curves, the TP region) and farther up the road ahead (Underwood et al., 1999; Lehtonen, Lappi, Kotkanen, & Summala, 2013). This time sharing between immediate and more anticipatory control seems to require executive attention (Lehtonen et al., 2012; see previous point about working memory use). This pattern may be partly due to the cognitive cost of planning the action several steps ahead, but probably also reflects the randomness inherent in the task environment: Unobservable properties of the situation make it difficult or even impossible to plan even relatively short action sequences in such detail that it would be possible to perform them in a “ballistic” fashion without feedback (Macuga, Beall, Kelly, Smith, & Loomis, 2007; Wallis, Chatziastrous, Tresilian, & Tomasevic, 2007).

7. Integration of visual or spatial information occurs across saccades, and memory representations can be used to orient in three dimensions. When working in a familiar kitchen, one immediately, without search, orients in the right direction when a particular utensil is requested. Land and Tatler (2001) discuss the visual strategy of a racing driver in terms of a rich memory representation of the lap, and the possibility of orient gaze and the vehicle in a way that takes into account road geometry beyond the range currently in view. Cavallo, Brun-Dei, Lava, and Neboit (1988) found in an occlusion experiment that when vision was occluded 2 s before entering a curve, normal drivers also were able to execute the appropriate magnitude of steering wheel rotation at the appropriate time, perhaps indicating the use of a “visual buffer” (Land & Furneaux, 1997). Look-ahead fixations in driving (point 6) also suggest that some kind of intersaccadic short-term memory representation used for trajectory planning may be constructed even in everyday driving (see discussion in Lehtonen et al., 2014).

Visual strategies in curve driving

For 20 years, “steering by the tangent point” has been the prominent theoretical model of visual
behavior in curve driving. Studies on car drivers’ gaze behavior during curve negotiation has shown that—with remarkable systematicity and stereotypicity—when approaching and turning into a bend drivers spontaneously direct their gaze toward the curve apex (for the definitions of apex and the TP used here, refer to Figure 1A). In the visual field this corresponds to the area surrounding the TP—the point where the visual orientation of the lane edge is reversed. In the visual science literature this behavior has come to be called “tangent point orientation,” and in textbooks “steering by the tangent point” is often presented as the default interpretation of the observed visual behavior.

TP orientation was empirically demonstrated and placed into a context of vision-action strategies1 in a seminal study by Land and Lee (1994). They were also the first to introduce the TP concept to the visual science community (but see Raviv & Herman, 1991, for a geometrical analysis in which most TP strategies reviewed here are already explicitly discussed). TP orientation here means that a substantial number of fixations fall within a region of the visual field spanning only a few degrees around the TP. Note that this definition refers simply to the observed behavior and does not imply that the drivers are actively looking at the TP (i.e., that the TP is the visual target of foveal gaze) or steering by the TP (i.e., that the VD of the TP, or local visual information picked up at the TP, is used to guide steering input).

A functional interpretation that the TP is used for steering clearly goes beyond gaze position data. While TP orientation has been replicated in many on-road studies (Underwood et al., 1999; Land & Tatler, 2001; Chattington et al., 2007; Kandil, Rotter, & Lappe, 2009, 2010; Lappi, Lehtonen, Pekkanen, & Itkonen, 2013a) and simulator studies (Marple-Horvat et al., 2005; Coutton-Jean et al., 2009; Authié & Mestre, 2011; Mars & Navarro, 2012), it remains contentious (a) whether it really is the TP itself that the drivers are looking at during TP orientation and (b) whether or not the TP is actually used for controlling steering.

TP models versus FP models

Land and Lee (1994) proposed that TP orientation is observed because drivers are looking at the TP (rather than some nearby points of interest in addition to or instead of the TP) and steering by the TP. Generally speaking, TP models such as this share the following assumptions: (a) The TP is the gaze target of foveal vision and the driver is “fixating” the TP rather than some other nearby point, and (b) the TP is tracked in order to pick up preview information of road geometry, used in adjusting steering in a visual steering strategy.

FP models are based on the assumption that the drivers visually track target points on the road surface the driver wishes his locomotor trajectory to fall on (i.e., on their FP). Thus, FP models posit that (a) a target point on the FP is tracked and (b) the FP target point is tracked in order to pick up preview information of road geometry used in adjusting steering in a visual steering strategy. Because these points are generally quite near the TP, this may still lead to TP orientation. In what follows, the discussion of TP versus FP steering models is organized by further classifying the models by whether they posit that the driver steers using the horizontal VD of focal steering point(s) or more global properties of the ambient optic flow (Table 1).

Another important (but rarely emphasized) distinction is whether the gaze targets on the FP are assumed to be travel points—which move with the observer in the three-dimensional scene (allocentric) frame of reference but may remain stationary in the observer’s (egocentric) frame of reference such as the forward-looking visual field—or way points—which are stationary in the three-dimensional scene but may move in the egocentric frame of reference. Examples of travel points are the TP or reference points on the FP referenced to the TP, such as FP reference point adjacent to the TP (FP_A, at the same vertical pitch as TP; Boer, 1996; Figure 1A) or beyond the TP (FP_B, at the same egocentric VD as TP; Wann & Land, 2000), but also points on the FP at a constant distance or a constant time headway ahead. A waypoint could be any point on the visible path, such as a marking on the pavement, or, if the driver plans his FP, a point such as the turn point (where the driver will turn the wheel entering the bend) or reversal point (where steering lock is reduced for exiting the bend) could be behaviorally salient as these are locations of changing trajectory curvature where the driver has to initiate some steering action—but they need not in principle present any distinctive visual feature.

Steering point (VD) models based on the TP and FP targets

The original explanation (Land & Lee, 1994) for TP orientation is that the TP is used as a steering point (Figure 2A through D). This term is here given a specific sense: a point in the visual field (a travel point) or a location in the physical scene (a waypoint) that the driver directly uses as a reference point for steering, where direct use means observing the VD^2 or depth distance of this point and using this information to determine the appropriate steering angle in a way that can be stated as a simple control law. (This is in contrast to using the observations to build a complex mental representation of the scene and using this to control
steering.) Most TP models treat the TP as a steering point, with the exception of the visual flow model explained in "Visual flow models." (Flow models are based on registering the flow pattern in a larger region of the visual field than a single point.)

Land and Lee (1994) actually put forward several models. One model proposed that (a) drivers fixate the TP (b) use the VD of the TP relative to the locomotor axis (Land & Lee, 1994; Land, 1998; Wann & Land, 2000) (c) set point control, maintaining a constant VDTP or keeping dVDTP/dt = 0 (Land & Lee, 1994; Land, 1998; Wann & Land, 2000)

Three-level set point control, tangent point VD as far point error term (Land, 1998; Salvucci & Gray, 2004)

Boer (1996) developed the steering point model based on a steering point on the FP. It is essential that the target point is on the FP, but, inspired by the TP result, he suggests a target "next to the TP but slightly into the road" (WPA) is used. It is not, however, essential that the target point is chosen in relation to the TP. Other waypoints such as WP, or a point chosen at a specific distance or time headway, could be used. Control is based on maximizing path radius (or minimizing the maximum lateral acceleration required to negotiate the turn at a given speed) within the geometric constraints of the lane edges. The model assumes that the driver executes a trajectory that takes them to a point on the FP, as follows:

1. The driver observes the VD of an FP target point relative to current heading, VDWP, and the time headway to the target THWP
2. Steer to produce linear RF when tracking a waypoint on the future path (a WPFP) (Wann & Swapp, 2000)
3. Two-level set point control, future path travel point (e.g., FPB) as far point error term (Salvucci & Gray, 2004)
4. Steering to produce a constant-rate inward sweep of a waypoint on the future path (WPB) (Wann & Land, 2000; Wilkie & Wann, 2003; Wilkie et al., 2008)
5. Planning and executing a minimum lateral acceleration trajectory to waypoint WPB (Boer, 1996)

Boer (1996) developed the steering point model based on a steering point on the FP. It is essential that the target point is on the FP, but, inspired by the TP result, he suggests a target “next to the TP but slightly into the road” (WPA) is used. It is not, however, essential that the target point is chosen in relation to the TP. Other waypoints such as WP, or a point chosen at a specific distance or time headway, could be used. Control is based on maximizing path radius (or minimizing the maximum lateral acceleration required to negotiate the turn at a given speed) within the geometric constraints of the lane edges. The model assumes that the driver executes a trajectory that takes them to a point on the FP, as follows:

1. The driver observes the VD of an FP target point relative to current heading, VDWP, and the time headway to the target THWP
2. Steer to produce linear RF when tracking a waypoint on the future path (a WPFP) (Wann & Swapp, 2000)
3. Two-level set point control, future path travel point (e.g., FPB) as far point error term (Salvucci & Gray, 2004)
4. Steering to produce a constant-rate inward sweep of a waypoint on the future path (WPB) (Wann & Land, 2000; Wilkie & Wann, 2003; Wilkie et al., 2008)
5. Planning and executing a minimum lateral acceleration trajectory to waypoint WPB (Boer, 1996)

Table 1. Steering models classified by gaze target (TP vs. FP) and stimulus characteristics (focal steering points versus ambient visual flow). VD = visual direction (horizontal visual angle relative to the "straight ahead" locomotor reference direction; z-axis). Note that this could be body axis, vehicle axis, or heading (which often but not always coincide). xL = lateral distance from the lane edge. Note that retinal flow is a special case of visual flow, but the abbreviation VF = visual flow in a forward-looking visual field (orthogonal to the z-axis). Other abbreviations refer to reference points explained in Figures 1 through 3. RF = retinal flow.

<table>
<thead>
<tr>
<th>Steering point (VD) models</th>
<th>Future path models</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Estimate path curvature (and steering angle) from VDTP and distance from lane edge xLE (Land &amp; Lee, 1994)</td>
<td>1. Estimate path curvature (and steering angle) from visual direction and distance of travel point FPB (i.e., VDTPB and DTPB) (Wann &amp; Land, 2000)</td>
</tr>
<tr>
<td>2. Set point control, maintaining a constant VDTP or keeping dVDTP/dt = 0 (Land &amp; Lee, 1994; Land, 1998; Wann &amp; Land, 2000)</td>
<td>2. Set point control, maintaining a constant travel point VDTP or keeping dVDTPB/dt = 0 (Wann &amp; Land, 2000; Wann &amp; Wilkie, 2004)</td>
</tr>
<tr>
<td>3. Two-level set point control, tangent point VD as far point error term (Land, 1998; Salvucci &amp; Gray, 2004)</td>
<td>3. Two-level set point control, future path travel point (e.g., FPB) as far point error term (Salvucci &amp; Gray, 2004)</td>
</tr>
</tbody>
</table>

Table 1. Steering models classified by gaze target (TP vs. FP) and stimulus characteristics (focal steering points versus ambient visual flow). VD = visual direction (horizontal visual angle relative to the “straight ahead” locomotor reference direction; z-axis). Note that this could be body axis, vehicle axis, or heading (which often but not always coincide). xL = lateral distance from the lane edge. Note that retinal flow is a special case of visual flow, but the abbreviation VF = visual flow in a forward-looking visual field (orthogonal to the z-axis). Other abbreviations refer to reference points explained in Figures 1 through 3. RF = retinal flow.
Figure 2. Steering points with the movement of the point indicated by an arrow. Steering points are depicted in both the visual field (first-person perspective view, egocentric movement) and the three-dimensional scene (bird’s-eye view, allocentric movement). (A) The TP on the lane edge. Maintaining $V_D_{TP}$ at a target value can be used in a control law to control steering angle or together with
though one must remember acceleration also depends on speed); and (c) the trajectory remains at all times within lane boundaries (for both left and right lane edges, \( |x_{LE}| > 0 \) at all points on the path).

Wann and Land (2000; see also Wilkie & Wann, 2003; Wilkie et al., 2008) have also presented models based on a FP target point beyond the TP as the steering point (FPB or WPB). A curvature estimation strategy suggested by Wann and Land (2000) is that the driver “looks through the bend” and determines steering from estimated path curvature \( \rho_{path} = 1/R = 2 \sin VDFPB/D_{FPB} \) (assuming a circular trajectory from current position to the target point, note that distance \( D_{FPB} \) must be estimated). A control strategy similar to the constant VDTP strategy above can also use a FP target (figure 1 in Wann & Land, 2000) wherein the TP is simply replaced by a FP travel point in the same VD (FPB). Then, steering control is based on maintaining \( dVDFPB/dt = 0 \) and \( VDFPB = x \), producing a (locally) circular path with a constant lane position \( x_{LE} \). Note that this will work only for a FP travel point, where the point of regard remains stable in the visual flow (vertically above the TP), while the point of fixation glides along the road. For an allocentric waypoint WPB (figure 2 in Wann & Land, 2000) the constant VD strategy would produce a logarithmic spiral trajectory overshooting the waypoint.

Another strategy (Wann & Land, 2000; see also Wilkie & Wann, 2003; Wilkie et al., 2008) is that the driver may steer so that a FP waypoint WPB sweeps from its initial offset \( VD_{WPB} = VD_{TP} \) to directly in front of the locomotor axis at a constant rate (which is half the vehicle rate of rotation), \( dVDFPB/dt = 1/2 (d\phi/dt) = constant > 0 \), and \( d^2VDFPB/dt^2 = 0 \) (for geometrical analysis, see also Wann & Swapp, 2000). The perceptual control variable here is visual acceleration of the point of regard, a variable to which humans are sensitive, especially if visual pursuit of the target can be performed (see Werkhoven, Snippe, & Toet, 1992).

Two-level steering point models

The TP has also been incorporated into the most influential theoretical framework for the visual control of steering: the approach based on the two-level control-theoretic model (Donges, 1978; see also, e.g., McRuer, Wade Allen, Weir, & Klein, 1977). Two-level control distinguishes between stabilizing control and guidance control, which in the visual science literature have been associated with different visual targets (Land, 1998; Salvucci & Gray, 2004; see also Land & Horwood, 1995; Figure 2E, F).

These models are an extension of earlier “closed-loop” feedback models based on canceling deviations from desired lane position. Some of the steering point models (e.g., the constant VDTP rule in Land & Lee, 1994) produce constant radius curvature at a constant lane position, while in the FP models there can be more flexibility in representing where the driver desires to place the FP target. Note that in psychological experiments the path is often assumed or instructed to follow the center of the lane. This is inconsistent with the actual trajectories people take (Spacek, 2005) and should not be considered a priori as an “optimal” trajectory without first determining the cost function drivers try to optimize the behavior to.

Once a desired lane position is determined, lane position error can be specified visually by the VD of a near point on the FP immediately in front of the vehicle (NPFP) or by a near point on the lane edge (NPLE), in which case the target VD depends on lane width and the depth distance \( D \) to the near point. If the driver wishes to maintain this lane position, then steering into the deviation of \( VD_{NP} \) from the target value will tend to cancel the error. This is stabilizing control.

Two-level models are motivated by feedback delays in the steering response (about 0.5 s—not to be confused with the time headway from current location to the point of fixation), which means that at high speeds steering cannot be achieved by stabilizing control alone. Using a near point as a set point in a delayed-feedback loop would lead to overcorrection and oscillation.

The two-level steering point models of Land (1998) and Salvucci and Gray (2004) assume that this guidance level control signal is based on monitoring the VD of a far point farther ahead (typically 1–2 s). Independent stabilizing control signals (VDNEAR as the error signal) and guidance control signals (VD_FAR) are then
weighted and combined in determining the steering input. “Anticipatory” information from a steering point farther up the road thus enables road curvature farther ahead to be anticipated, thus creating smoother steering. Appropriately adjusted time headway and lag are needed to avoid oversteering into the bend. (Note that the control law is still a feedback loop—the anticipatory information is not derived from prediction, open loop, but specified by visual information.)

Salvucci and Gray (2004) called areas in the visual field where the stabilizing and guidance level steering points lie the near zone and the far zone, respectively. The role of eye movements directed to targets in the far zone or the near zone as needed can be conceived as serving the needs of control processes at these different “levels” of control. The near zone, however, is usually assumed to be peripherally monitored (see Summala, Nieminen, & Punto, 1996), in which case gaze will be predominantly directed toward the guidance level steering point(s).

This connects the two-level models to TP orientation because the TP is considered one potential guidance level target point and TP orientation can be interpreted as fixating a far point used in two-level control based on two steering points (Land, 1998; Salvucci & Gray, 2004). The far point, however, can be any “salient distant point with which the model can monitor lateral stability and, given its distance, maintain a predictive steering angle that compensated for the upcoming road profile” (Salvucci & Gray, 2004, p. 1236). The TP and the FPB are illustrated in Figure 2E and F as examples. But as Salvucci and Gray (2004) emphasize, the model will work with any points in the visual scene that behave appropriately (i.e., are cross-correlated with required steering input with some constant or known delay).

There are two further points worth noting. First, the two point models make no essential use of properties that differentiate the TP from points on the FP. Second, while a minimum of two points are necessary for two-level control (and seem to be sufficient to produce stable behavior under some simulated steering tasks), there is no apparent reason why the driver could not in principle use both near points on the lane edge and road ahead for stabilizing control as well as both the TP and a FP travel point in the far zone.

The two point control scheme cannot be cast as a TP model or an FP model in simple either/or terms: In the Salvucci and Gray (2004) model, the driver may use the TP as a far point, in which case the model can be considered a “TP model,” but as they note, the model could just as well use a FP target point in the far zone (p. 1236). In this case, the model becomes a two-level “FP model.” There are numerous points that could act as the FP travel point in the far zone (FP_A, FP_B, a travel point at a constant distance or a constant time headway, or in closed bends with very short sight distances, even the occlusion point; Lappi et al., 2013a). Indeed, as more than two points could be used for control, the term two-level model gives perhaps an overly restricted view of how steering points might be used in visual control of steering (see discussion in “Conclusions: The next 20 years”). The idea of simultaneously tracking the visual motion of multiple points in the scene brings us to the discussion of visual flow.

**Visual flow models**

Optic flow refers to the relative angular movement of visual pattern in a textured scene caused by observer motion (Figure 1B). The origin of the concept is Gibson’s (1958, 1986) ecological optics, and control laws for steering based on optic flow. The concept itself, however, is not restricted to Gibson’s more general (and more controversial) theory of ecological optics and direct perception (i.e., that vision should be considered as the use of information available in the optic array rather than the production of representations by means of possibly complex computational operations, such as multiple coordinate transformations), and it has been applied in artificial intelligence and mobile robotics (for an overview, see Wilkie, Wann, & Allison, 2011).

Optic flow can be represented as a vector flow field of visual texture (pattern, features) in an image onto which the scene is projected (a visual field, which can either be a purely abstract image, or a physical image surface, such as the windscreen, the photocell of a camera, or the retina). Often optic flow is displayed on a plane projection perpendicular to the direction of travel (locomotor axis). Here, this projection is referred to as visual flow. Retinal flow is the projection of optic flow on the surface of the retina. This is the physical visual stimulus available to the brain.

The simplest visual flow patterns are caused by rectilinear motion directly toward or away from a textured wall (Figure 3A, inset A1). When moving toward the wall, the flow pattern is a uniform radially symmetric expansion from a focus of expansion. Movement away from the wall will induce a pattern of uniform contraction. In a more complex scene or when the trajectory is more complex, the optic flow field structured by the environment becomes more complex as well. Figure 1B schematically illustrates optic flow in a road scene when turning into a bend. Here, translation and rotation occur together and there are multiple objects at different distances. Note that there is no globally coherent flow, but there are regions of locally coherent flow.

In addition to the complexity of visual flow, when we move our head or eyes this further affects retinal flow. As the visual axis will typically not be aligned with the
locomotor axis (but anticipating the direction change) and as the eye itself may also rotate during locomotion, the relationship between optic flow and retinal flow can be quite complicated (Regan & Beverley, 1982; Kim & Turvey, 1999; Li & Warren, 2000; Wann & Swapp, 2000; Authié & Mestre, 2012). This appears to create a problem for models that assume steering is based on a visual analysis of optic flow: When all the brain has to go by is retinal flow (and extraretinal information such as oculomotor corollary discharge and vestibular

Figure 3. Geometrical relationships between the TP, optic flow, and path geometry. (A) A constant curvature path in a constant curvature bend (no steering change required). The FP is on an equal flow cylinder. The TP is on another equal flow cylinder (zero horizontal flow; visual flow is vertical). Multiple TPs (where line of sight is tangential to a curve following the road curvature) are present on the zero flow cylinder (see Figure 1B). Inset A1: Optic flow during linear translation; a focus of expansion specifies the current heading. Inset A2: Optic flow during rotation; direction and magnitude of flow specify direction and rate of rotation. Inset A3: More complex flow is generated by simultaneous translation and rotation (of locomotor axis, the reference axis for the visual field). Here, equal flow cylinders emerge (green, red, blue), with points on the FP (blue) having equal (negative) flow and TPs (red) having zero flow. Note that this is a property of the visual flow field and is not dependent on demarcated lane edges around the path, nor is the TP on the lane edge unique. Eye movements will determine how this pattern appears in retinal flow. Inset A4: In a visual flow representation (see Figure 1B), waypoints on an equal (negative) flow cylinder have the same horizontal component as do TPs on the zero flow cylinder. (B) Curve radius decreases. Constant radius path (no steering change) would result in lane departure on the inside edge (oversteering). TP lies on a constant flow cylinder with inward flow. (C) Curve radius increases. Constant radius path (no steering change) would result in lane departure on the outside edge (understeering). TP lies on a constant flow cylinder with outward flow. Note that in B and C the movement of the TP is in the direction of flow, which is opposite to steering error (i.e., the direction of flow at TP indicates direction of the required correction).
information), how can it recover optic flow? The problem is circumvented, however, in the active vision approach where the retinal flow pattern itself is used to control both eye position and locomotor steering—without the need to recover visual or optic flow by subtracting the effects of eye movements.

Kim and Turvey (1999) presented an analysis of retinal flow, whereby retinal flow lines can specify the curvilinear path in the retinal flow field analogously to the way the focus of expansion specifies heading in the optic flow field during linear translation.

It seems intuitively a rather remarkable property of optic flow that if the driver is on a (circular) path that will take him or her to the point of fixation—conversely, if the driver is looking at a point where his path must have zero flow.

Wann and Swapp (2000) independently presented a model based on an idea of using this optic property in an active gaze/retinal flow strategy for steering. Here, the eye movements themselves are essential: The driver fixates and tracks a waypoint on the road edge (a salient feature), and because horizontal flow is equal to zero flow cylinder the entire path must have zero flow.

A TP model based on an analysis of optic flow has also been put forward. It is based on the fact that when the trajectory curvature of the observer matches road curvature, visual flow at the TP is vertical (i.e., the TP is on the zero flow cylinder; Raviv & Herman, 1991). Horizontal optic flow at the TP can therefore be used as information about steering error (see Figure B, C). Inward flow indicates understeering (or imminent increase in road curvature); outward flow indicates oversteering (or reduction in road curvature). When the local curvature of the path matches the road curvature ahead, then the TP on the lane edge falls on a zero flow circle. The most recent formulation is by Authié and Mestre (2012), who are the first to argue for the model based on behavioral data, but the control scheme can be found in Raviv and Herman (1991), and Land and Furneaux (1997) also discuss this property of flow at the TP.

This analysis actually holds not just for the TP on the lane edge but for all TPs where line of sight at that point is tangential to a curve following the road curvature (see Figure 1A and Figure 3A, main picture and insets A3 and A4). This more general concept of a TP is not dependent on the presence of “lanes” demarcated by painted edge lines. In driving studies the TP on the lane edge figures prominently, and from the point of view of the driver’s task of maintaining the car between lane boundaries it may be considered particularly relevant. To say that the TP is a “singular” point for this strategy is not quite accurate, however. Authié and Mestre (2012) suggest that the TP on the lane edge is an “optimal” gaze target for this strategy because it is a visually salient point (indeed, during steady-state cornering the inside road edge presents a stable visual feature), and because horizontal flow there is at a minimum it has been suggested that the eye can “rest” there without being “dragged” around by an optokinetic reflex following the flow (see Land & Furneaux, 1997, and below). According to the model, the driver (a) fixates the TP on the road edge (a salient feature), (b) observes regional visual flow around the point of regard, and (c) adjusts steering in the direction of the horizontal component indicating steering error or change in curvature.

Although based on a sophisticated geometrical analysis, in practice this model is essentially similar to the steering point models based on canceling target point VD movement (see Table 1), as the horizontal drift of the TP corresponds to the flow at that time—only here the stimulus parameter registered by the visual system to determine the horizontal movement is assumed to the deviation of visual flow relative to the allocentric vertical, not the egocentric VD of a steering point.

On-road studies of gaze behavior during curve negotiation

The reviewed models are not only geometrical exercises in analyzing curve geometry and the visual
field projection (stimulus parameters) but also have been put forward as hypothetical models of how people driving cars (and in other similar forms of locomotion) may actually behave. However, with the exception of Land and Lee (1994) where gaze data are presented from two drivers, the papers presenting the models do not present experimental driving data. Next, the relevant on-road experiments are reviewed in order to assess how well the different models account for the data collected during the past 20 years, in experiments inspired by the Land and Lee (1994) TP paper.

The critical question is how well real-world data can constrain model selection. With so many models—all of which make the same qualitative prediction of TP orientation, remember—settling the issue of empirical adequacy of the models is going to be more complicated than simply recording “where people look” or identifying various points in the visual field and trying to determine whether the point of regard is “in that area of interest (AOI).”

Classical TP AOI results

In their seminal study, Land and Lee (1994) found the following: (a) The TP was likely to be fixated immediately before turn-in (1–2 s before the car “entered the bend,” defined as zero crossing of steering); (b) 0 to 1 s after entering into the bend the driver’s gaze was within 3° of the TP for more than 75% of the time; (c) 0 to 2 s after entering the bend gaze was within a 3° AOI for about 50% of the time; and (d) in the distribution of fixations, the area of highest fixation density was within 1° of the TP.

The AOI sizes and gaze catch percentages in follow-up studies have varied somewhat (for details on the methods and analyses, see Supplementary Tables S1–S9). The complete list of findings reads as follows: Underwood et al. (1999) reported dwell times within 2° of the TP for experienced drivers (13% in closed curves and 12% in open curves) and for novice drivers (11% in closed curves and 14% in open curves; all differences between the groups was not significant). The results for the 6° AOI have a different pattern—but this AOI could be so large that much of the road and scenery falls within it, and one cannot draw firm conclusions (see Figure 4 and the section on AOI overlap below).

Kandil et al. (2010) report 67.3% ± 2.1% (mean ± SEM) fixation times within 2° of a TP or lane markings AOI in the beginning of a turn and 54.5% ± 4.0% in the later parts of the turn. Lappi and Lehtonen (2012) report 52% ± 23% to 72% ± 24% (mean ± SD) fixation times within 3° AOI TP orientation in different curves during curve approach (prior to turn-in) and 41 ± 28% to 67 ± 22% during curve entry (after turn-in). Lappi et al. (2013a) reported 39% (left) and 48% (right) gaze catch in 3° AOI at the CL (left) and TP (right) during curve approach and 26% (left) and 27% (right) during curve entry. These studies were all conducted on rural roads. In two studies on motorway ramps, Kandil et al. (2009) reported “looking at the TP” as much as 75% of the time, whereas Lappi et al. (2013b) reported a mere 23% of gaze in a 3° AOI (35% in 4°, 56% in 6°) when entering an on-ramp.
The pattern and causes of variability are not currently well understood. Variability in the results is likely due in part to different methods of measurement and analysis and in part to differences in behavior. Factors such as road geometry, the phase of the curve (approach, entry, cornering, exit), potential individual differences in drivers, or driving speed could all play a part.

FP AOI results

There are more technical challenges in representing the FP than the TP in an egocentric (gaze angle) coordinate system with real-world data. This is probably one reason why the FP as a potential gaze target has been less explored in most on-road experiments, and steering by the TP has been the default interpretation in on-road studies—even though the majority of the steering models predicting TP AOI orientation (Table 1) are actually FP models!

What is required to assess FP AOI orientation is an estimate of the angular positions of the points in the visual field corresponding to FP. The TP can be clearly identified (manually or algorithmically) from an image of the road scene because it is a salient geometrical visual feature (requiring only a mapping from the image coordinates to gaze angles). This is generally not the case for the FP, which is a more complex geometrical object and as such is not visible in a forward-looking image (it is, after all, the future path). Also, while the TP is indeed a point, the FP is a curve (whose length varies, even within a bend, depending on the road geometry and sightlines). This makes comparing the models’ predictions with AOI methods problematic because a larger AOI is—of course—expected to get more hits even by chance. Conversely, if you were to dynamically scale the TP AOI size to correspond to the FP AOI size, the AOIs risk becoming so big that they overlap completely (the AOI overlap problem; see below).

Selecting a singular reference point on the FP is not entirely straightforward, either. For one, unlike the road edge TP, the FP is not associated with any distinctive image features to differentiate it from adjacent parts on the road ahead. The exact location of the “true” reference point(s) is of course selected by the driver and may depend on road geometry or individual differences and is not known a priori. Most of the FP models are not actually very specific on where on the FP the reference point should be found. The exact placement in depth is not essential in the models. Boer (1996) and Wann and Land (2000) propose target points adjacent and beyond the TP, respectively. Empirically, gaze distribution on the road generally appears to be farther in the far zone, above and beyond the TP (Lappi et al., 2013a, 2013b; see also figure 2 in Land & Lee, 1994), favoring FP reference points beyond the TP (FPB, WPB) over FP reference points adjacent to the TP (FPA, WPA). But even in these models there is no fundamental reason why the TP (rather than, say, time-distance) should be the true reference. Salvucci and Gray (2004) discuss their two-level model in terms of TP orientation (but in the actual implementation they use far points on the FP 1°–9° below the horizon).

Lappi et al. (2013a) evaluated the TP against FP targets (and the suitability of AOI methods for this type of analysis) by using a parametric representation of the visual projection of the FP. This study was also the first to assess overlap between the TP AOI and areas of interest placed on points on the FP, previously raised in discussing potential problems for interpreting AOI data on TP orientation (in Robertshaw & Wilkie, 2008).

The main conclusion, drawn from the analysis of AOI overlap, is that—especially during curve approach—the proximity of the target points renders the traditional AOI approach quite ineffective as far as identifying different gaze targets is concerned. Lappi et al. (2013a) also present a slightly different method of computing gaze catch percentages, where each gaze observation is clustered to the nearest (a priori defined) reference point. This method is in some ways preferable to the traditional way of counting a gaze observation as an AOI “hit” whenever it falls within a predefined distance from the reference point regardless of whether other reference points are nearby. For one, it does not require an arbitrary AOI size to be set. It does not, however, solve the fundamental problem of reference point contiguity because misclassification may still occur—especially if spurious AOI reference points are used, which the driver does not in fact ever target! Thus, beyond saying that the gaze distribution tends to be “in the far zone” and more beyond than adjacent to the TP, AOI-based gaze distribution results in on-road studies seem to add little.

Kandil et al. (2009) attempted to differentiate between the TP and FP by means of an experimental manipulation using different gaze instructions. The participants were first instructed to drive naturally but then, in a second phase of the experiment, they were instructed to either look at the TP (to maintain permanent fixation on TP) or use a gaze sampling waypoint strategy (i.e., successively look for and keep fixating for several seconds at points on the FP of the car). That study reports smoother driving in the TP condition compared with the gaze sampling condition (less variability in the steering signal). This the authors interpreted as evidence for the TP hypothesis and against steering models based on targeting the FP. However, as the authors themselves note, the normal pattern in driving is not continuously “staring” at a single point, such as the TP. Also, the “gaze sampling” instruction required a highly artificial kind of FP orientation. Fixations tracking the same
target location for several seconds do not reflect the normal fixation pattern in driving (Lappi et al., 2013b; Lappi & Lehtonen, 2013), where the tracking fixations on the road (i.e., optokinetic pursuit movements; see below) last for only a few hundred milliseconds. The authors state that in their raw data they did not see evidence for gaze sampling (tracking fixations), but this is hardly conclusive—especially if by gaze sampling they mean the kind of artificial strategy instructed in the gaze sampling condition. The potential confound of AOI overlap is not discussed in the study, and a quantitative estimate of the projection of the FP in the visual field was not used. Instead, the gaze target was determined by visual inspection alone, and the classification (“TP” or “road”) was then coded manually into the data. In summary, to what extent FP orientation does or does not occur in normal driving is by no means conclusively established by the Kandil et al. (2009) results.

**Difficulties in the interpretation of AOI data**

Area of interest methods rely on quantifying the relative frequency of gaze catch in an AOI, centered on a putative target point. In on-road studies this is usually the TP, which is easy to identify in a forward-looking camera image making TP orientation easy to observe, but AOI methods are poorly suited to differentiating between TP models and FP models. Data in support of one model (TP orientation) in naturalistic driving are not automatically data against other (FP) models.

Because on-road studies lack experimental control of where to “put” the reference points, all the gaze targets are usually quite close together (see Figure 4). With realistic AOI sizes and typical curve geometry, the AOIs frequently overlap (Lappi et al., 2013a). This means that an observation of gaze position in one target point’s AOI (e.g., TP AOI) would be an observation of gaze position in another target point’s AOI as well (e.g., FP_B AOI). Unless both are included in the analysis, the author may conclude that looking at the TP is supported, but what cannot be determined is to what extent it is supported over the relevant alternative hypotheses. Thus, FP_B could be equally or even more “supported” by the same data! AOI overlap creates an experimental confound, which was rarely addressed in the early TP studies, and a functional interpretation of the TP as a steering point cannot be inferred from a “high” percentage gaze catch by a suitably “small” TP AOI. All “classical” TP orientation results based only on AOI gaze catch percentages in a single AOI are ambiguous, with respect to differentiating between TP models and FP models.

Perhaps even more insidious is that the close vicinity of the potential targets can lead to false positive results, identification of FP (or TP) orientation when the point around which an AOI is placed is not in fact looked at. This is clearly unacceptable. If there were only one theoretically possible steering point, then any measurement error would lead to a conservative weakening of the result. However, with the visual field littered with potential targets, spurious hits to AOIs are bound to occur. Making the AOIs very, very small and striving for perfect accuracy is no panacea either, as a 3° AOI is already getting close to the size of the fovea, and if steering employs parafoveal and peripheral (e.g., flow) information, then where exactly the fovea is directed in naturalistic steering tasks may not even be the right question to ask. Overall, simple AOI methods are poorly suited to provide data simultaneously on the different gaze targets of the different models.

**Beyond AOI methods**

During curve negotiation, the curve naturally “opens up” as the TP is displaced into a more eccentric position relative to the locomotor axis (see Lappi & Lehtonen, 2012, and the appendix to Lappi et al., 2013a), and the FP in the far zone typically also moves vertically farther from the TP. However, only in very wide curves and with extremely high spatial resolution can one hope to resolve between the TP and even a small number of FP AOIs. In the study by Lappi et al. (2013a), a different means of distinguishing between TP and FP models’ predictions was used: It was shown that when the visible road in the far zone subtends a larger vertical angle, the gaze distribution (within a sector in the VD of the TP) is also displaced vertically from the TP. This dependency between the vertical angular subtension of the road and the vertical distribution of “TP oriented” gaze suggests that at least some of the fixations in the TP direction are in fact directed towards the road surface. Investigating the vertical distribution of gaze in the visual direction of the TP thus provides stronger (and complementary) evidence that future path orientation occurs in real driving, than what one gets by simply calculating the gaze percentages in AOI’s on the future path.

In another study by Lappi et al. (2013b), optokinetic pursuit movements6 were used as another means to search for evidence for or against FP orientation (Figure 5). Optokinetic pursuit movements are slow eye movements that follow a moving target object, track a location moving in relation to the observer (smooth pursuit), or follow local optic flow by reducing retinal slip (optokinetic reflex). In optokinetic nystagmus (OKN) the slow phase optokinetic movement is followed by a saccadic movement that resets gaze. This creates the characteristic periodic pattern (nystagmus), which has been recently demonstrated to occur in simulated (Authié & Mestre, 2012) and real (Lappi & Lehtonen,
The “fixations” during curve driving and “TP orientation” are thus not fixed in the egocentric coordinate system. The eyes rotate horizontally. If the direction and rate of this rotation is consistent with the visual rotation of the parts of the road surface in the far zone beyond the TP—based on analyses reviewed above—this would support FP models. (A vertical downward rotation would be predicted by an optokinetic reflex during TP fixation and no rotation with perfect TP fix.)

Lappi and Lehtonen (2013) and Lappi et al. (2013b) presented data that indicated that in real driving optokinetic pursuit indeed has a horizontal component, in the direction opposite to the bend, and that the magnitude of this horizontal component was approximately equal to one half of the vehicle yaw rate. This OKN is both qualitatively and quantitatively consistent with coherent optic flow in the area where the gaze appears to land (i.e., far zone above and beyond the TP and spread along the road with a large horizontal variation). This qualitative pattern is also apparent, incidentally, in the original Land and Lee (1994) results, at least in the figures—Land and Lee (1994) do not discuss this aspect of their data. The reasoning, however, is at the present state somewhat indirect, based on the analysis of the optic flow field in Figure 3 (three-dimensional projection of optic flow has not yet been physically measured and modeled in any TP or FP study).

On balance, this pattern of results—both the distribution of gaze and gaze behavior during fixation—is consistent with the drivers’ targeting points on the road surface in the far zone beyond the TP instead of, or in addition to, the TP and is perhaps most consistent with the FP optic flow model of Wann and Swapp (2000) and the steering point model of Boer (1996), which predict tracking fixations on the FP.

It should be emphasized, however, that these results apply to the cornering phase (constant steering with minor corrections), whereas most data on TP orientation have been collected in curve entry or around the turn point. It is possible that different targets and different strategies may be involved in different task phases (i.e., the TP could be used in deciding when and by how much to turn the steering wheel, though one must be cautious because the AOI overlap problem is more pronounced in the early phases). After all, it is typical in naturalistic tasks that sequential organization of the task is closely coupled to the patterns of eye movement behavior.

Given the flexibility of human visuomotor behavior, any model that posit one “best” point in the visual scene (the TP or a steering point on the FP) seems less realistic than to assume multiple reference points that the driver can use as cues to determine appropriate speed and steering. If this is the case, the real question then is not whether the target is on the road edge or on the FP but rather what the targets are, when they are targeted, and what elementary actions that make up the skill of driving are used to control each of them.

Conclusions: The next 20 years

Since Gibson and Crooks (1938), driving has been a theoretical and empirical real-life test bed of visual
control of locomotion in a complex, three-dimensional environment. For example, in a recent review of 25 years of research into “how eye movements cope with real-world visual and cognitive demands” (Kowler, 2011), driving (“real roads, not simulators!” p. 1474) was identified as a core task, along with tasks such as reading and sports. This is true of steering behavior in particular, which here refers to the control of the locomotor axis of a vehicle (not to rotating the steering wheel as such; in a single track vehicle steering is achieved by countersteering and roll, and in recovering from a skid the steering wheel rotates in the direction opposite to vehicle yaw rotation).

The FP and TP models have been proposed as accounts of car driver visual and steering behavior, and this paper has focused on the theories that account for and are constrained by available data on car driving. Consequently, the empirical studies reviewed above all deal with steering a car (a) on a visually delineated road lane (with painted edgelines), (b) at everyday speeds, and (c) in a contemporary, developed Western country traffic environment. (The Land & Tatler, 2001, paper is an exception with respect to items b and c.)

To be of fundamental scientific interest, the mechanisms and strategies need, of course, be generalizable to other modes of locomotion and a wider range of tasks and environments. For while understanding driving has an applied interest as a time-constrained task in a hazardous environment to which hundreds of millions of people are exposed every day (i.e., the global traffic system), and a detailed understanding of steering behavior would undoubtedly contribute to practical applications in human performance and safety, of more immediate interest for the visual scientist is what we can learn of driving as a class of the visual control of locomotion. From this perspective, it is incidental that most driving humans do happens in the road transport system (but by no means all—e.g., in the military and in sports, driving typically occurs off road or on purpose-built tracks).

Many of the models are, indeed, quite general and make no special assumptions specific to cars on the road. Eye movements tracking FP waypoints with an optokinetic pursuit eye movement create a pattern where the FP (constant radius up to the point of fixation) is specified by vertical retinal flow (Kim & Turvey, 1999; Wann & Swapp, 2000). This is a purely geometrical result that does not depend on “a road” to be visually demarcated in the scene and is not specific to any particular mode of locomotion. In car driving experiments, there is an understandable tendency to identify target points on the FP with reference to the TP (adjacent to the TP as in Boer, 1996, or beyond the TP as in Wann & Land, 2000), but this is more to do with the saliency of the TP (or the ease with which it can be identified in a camera image for placing an AOI) and the historically dominant position of the Land and Lee (1994) paper rather than any deep theoretical reason. The TP itself is usually identified with the TP of the lane edge—but this concept is also more general (Raviv & Herman, 1991). On a constant radius path, TPs emerge as points of vertical flow in the visual flow field (and the retinal flow field if the TP is fixated, in which case points on the FP have outward retinal flow). No physical edgelines are required for the definition of a TP, just as no physical line need designate the FP the driver intends to follow. (It is interesting to observe, however, that whenever humans have constructed roads, it is always the edgelines that one must not cross that are visually designated by curbs, paint, and so on, and not the lane center.)

Some of the models (TP models 1 and 2 in Table 1) do seem to suggest a very simple strategy of identifying and fixating (i.e., maintaining the point of regard on) just this one visual feature in the visual field (i.e., reversal point of a painted edgeline) and using it to maintain a constant lane position. This is probably too simple to be realistic in terms of either the information represented in human path planning or the visual behavior and locomotor trajectory.

Starting with the locomotor trajectory, the normal track pattern is to cut to the inside of the bend (Spaceck, 2005), and a model that takes maintaining a constant (central) lane position as the normative or descriptive standard is undermined even before considering the visual behavior.

In on-road studies, in particular, there seems to be a tendency to focus on one “steering point” (usually the TP), and the results are sometimes presented (and even more often subsequently interpreted) as vindicating a particular steering point—even if the (AOI) data do not really strongly favor one target point or class of models over the others.

In fact, what is easily overlooked when presenting the issues in terms of dichotomies (TP or FP? A steering point or visual flow?) is that many of the TP and FP models are not mutually incompatible. Real-world steering may employ multiple visual targets in parallel (waypoints embedded in visual flow, travel points with VD), and while foveal gaze can be directed at only one point at a time, peripheral vision is constantly available, and the normal pattern is not to “stare” at any fixed point in the visual field but rather to “scan” the road scene (Underwood et al., 1999; Green, 2002; Kandil et al., 2010; Lehtonen et al., 2012, 2013, 2014; Lappi et al., 2013a), possibly to glean steering-related information from multiple targets. Salvucci and Gray (2004) are keen to demonstrate that two points can be sufficient (see Land & Horwood, 1995; but also see Cloete & Wallis, 2011, for a critique), and the TP models (Land & Lee, 1994) show that sometimes you can make do with only one. But that does not mean...
real-world control should not have built-in redundancy in terms of a larger number of reference points for more robust control (Sharp, Casanova, & Symonds, 2000; see Figure 6).

Even though only one travel point or waypoint can be foveated at a time, this does not mean that more are not tracked covertly—perhaps in a similar manner to the multiple object tracking paradigm (Pylyshyn & Strong, 1988; Pylyshyn, 1994, 2001). Interestingly, this is a task known to involve intraparietal cortical areas (Xu & Chun, 2009), a region of the cortex that is implicated in processing optic flow and path-related visual information and that plays a role in controlling eye movements (Billington et al., 2010; Field et al., 2007).

Multiple travel points weighted due to decreasing anticipatory cue value with increasing distance could be, for example, at fixed time distances from the observer, from current position to up to 2 to 8 s in the future. These types of models have been around since the 1960s in the vehicle engineering literature (see, e.g., Macadam, 2003). (B) Multiple waypoints on the FP (locations where the action sequence moves from one phase to the next). “Guiding fixation” on the outcome of the current phase and initiation of the next phase with occasional “look-ahead fixations” farther ahead facilitating path planning would be in line with the qualitative principles of gaze behavior in other naturalistic tasks (outlined in “Driving as real-world visual behavior”).

What visual cues are used, how they are represented, and how eye movements are used to sample them are among the outstanding questions in understanding the visual basis of vehicle-assisted locomotion. Despite more than 100 years of eye movement research (and more than 100 years of driving and psychological research on driver behavior), we are still only beginning to answer these questions. To what extent the visual strategies reviewed here apply to driving or generalize across modes of locomotion available to humans requires detailed study of steering behavior in both controlled and naturalistic contexts. This will, eventually, differentiate between general principles and task-specific cues and techniques, and here laboratory and field experiments have complementary rather than competing roles. Details of information processing strategies—and the underlying neural mechanisms—are usually best revealed by carefully controlled.
psychophysical, behavioral, and neurophysiological experiments. But whether the strategies (and, by implication, neural mechanisms) are actually used in the real world can be determined only by careful analyses of natural behavior in its ecological context.

One place where real-world experiments can be relevant to testing and developing theoretical models of driver behavior is the OKN pattern. It is predicted by waypoint models, but auxiliary assumptions about optokinetic reflex mechanisms are required to incorporate it into travel point models. So while the presence of OKN is not conclusive proof that waypoints are being foveated, it is as yet unclear how such periodic gaze behavior would emerge from strategies based on fixating travel points rather than tracking waypoints with a pursuit movement. One possibility would be that while the “top-down” designated far point (e.g., TP, FPB) itself does not move in the egocentric frame, the point of regard is “dragged” away from the target by an optokinetic reflex elicited by optic flow (Land & Furneaux, 1997; Authié & Mestre, 2011). This is followed by a catch-up saccade, which thus creates the observed pattern of OKN (for a more detailed rundown of the alternatives, see the appendix in Lappi et al., 2013b). As far as the visual steering strategy and oculomotor control are concerned, then, the driver is “trying” to fixate a travel point (rather than making a smooth pursuit movement tracking a waypoint), but this higher-level goal is overridden by a lower-level reflex aiming to reduce retinal slip (in some to-be-specified region around the point of regard). Thus, the growing body of real-world data will require in the future more precise hypotheses about the different levels of cognitive (and cortical) control of eye movements—both in locomotion generally and driving in particular. This is surely a good thing.

Precise control of stimulus parameters and behavior enables laboratory studies to better isolate mechanisms and establish causality, while field research is required to determine the strategies used in ecological settings and to validate simulator and laboratory results. The main difficulty with interpreting the results and judging the empirical impact of the many laboratory experiments and simulator studies on TP/FP orientation, for example, is that for firm conclusions to be made on the basis of a simulator finding, the assumption must be made that the behavior of interest is qualitatively or quantitatively similar in the simulator and in the real world—at the level of dependent variables or specific performance measures. Generally, this type of external validation is not available for any of the more sophisticated measures (that go beyond TP orientation).

Overall, the more straightforward replications of the basic TP orientation result (Marple-Horvat et al., 2005; Coutton-Jean et al., 2009; Authié & Mestre, 2011; Mars & Navarro, 2012) can be considered to be “validated” by on-road studies at a quantitative level. As for FP models, the OKN results (Lappi et al., 2013b; Lappi & Lehtonen, 2013) and results on look-ahead fixations (Lehtonen et al., 2013, 2014) can be considered to support the gaze polling results of Wilkie et al. (2008). Wilkie and colleagues have run a series of simulator experiments that have been steadily building up to a case for FP strategies and the use of multiple cues (including visual flow) in steering (e.g., Wilkie & Wann, 2003; Robertshaw & Wilkie, 2008; Wilkie, Wann, & Allison, 2008). For example, they have demonstrated that instructing participants to look at different parts in the virtual road scene affected the trajectory and that instructing participants to take different paths through a curve affected the point of fixation distribution. (When instructed to maintain a constant lane position at various distances from the lane edge, gaze was correspondingly distributed at different lateral positions in the lane: When instructed to cut the corner, the gaze was closer to the lane edge.) What is particularly commendable is that the results are discussed in terms of a steering model (centered on the active gaze fixation system) based on the relative weighing of multiple visual and nonvisual cues. On the other hand, it is not entirely straightforward to generalize results supporting a FP strategy in such simulator experiments—with explicit path instructions—to free driving or cycling on real roads.

It should be emphasized that this difficulty does not stem from inherent shortcomings in these simulator studies per se—they are in part a reflection of where the methodology of on-road experimentation needs to step up. On-road studies bring benefits of ecological validity, but usually at the cost of methodological challenges that make the data difficult to interpret. Thus, what one would hope to see in the next 20 years is development of on-road research and data-analysis techniques that go beyond simple summaries of AOI catch percentages and begin actually modeling the rich ecological three-dimensional stimulus. This would also serve to make on-road research more relevant to laboratory studies of visual behavior. Ideally, what one would like to see is real-world data analyzed in terms of real three-dimensional information about the scene (sightlines, distances) and simulator experiments run in a virtual three-dimensional environment constructed on the basis of the same physical measurements. This type of modeling of one particular class of visually guided locomotion in a complex, but pleasingly regular, three-dimensional environment—driving—should tie the study of such everyday behaviors even more tightly to our understanding of how the human brain achieves locomotor control: how it represents visual space and by what principles and strategies it organizes the eye...
and hand movements required for complex skills we so easily exhibit in our daily lives.

**Keywords:** locomotion, steering models, real driving, future path, tangent point

**Acknowledgments**

Incisive comments and criticisms by J. P. Wann, M. F. Land, R. M. Wilkie, and J. J. O. Pekkanen and E. P. Lehtonen on previous versions of this material are gratefully acknowledged.

Commercial relationships: none.

Corresponding author: Otto Lappi.

Email: otto.lappi@helsinki.fi.

Address: Cognitive Science Division/Traffic Research Unit, Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland.

**Footnotes**

1These are visual strategies that express oculomotor or locomotor control in terms of a simple rule or a control law that states a dependency between stimulus (three-dimensional scene layout, optic flow, visual field structure, or retinal image) and behavior. These types of models are particularly suited to ecological studies of visual behavior because model testing and development can proceed without detailed knowledge of the underlying mechanisms producing the sensorimotor transformations.

2Note that visual direction has here been used to refer to horizontal angular position of a gaze target in the forward-looking visual field—not the retinal visual field but one relative to the locomotor frame of reference. This could be the body axis, vehicle axis (determined visually in a car by a hood emblem, wings, or A-frame), or locomotor axis (i.e., instantaneous heading, which would not be directly visually specified during locomotion on a curved trajectory). If heading can be recovered from the retinal image or extraretinal information, then the VD relative to locomotor axis can be computed in retinocentric coordinates. If points on the vehicle frame are visible, then the VD relative to vehicle axis can be computed in retinocentric coordinates. And if the tangent point is fixated, then the VD relative to body axis can be recovered from extraretinal signals specifying the eye-in head and head-in body angles (Wann & Wilkie, 2004, pp. 378–379).

3\[ \rho_{\text{path}} = f(\theta_{\text{sw}}, v) \], which is based on the understeer/oversteer characteristics of the vehicle for the current speed \( v \). The task is to match path curvature \( \rho_{\text{path}} \) to bend curvature \( \rho \).


5They explain the pattern of higher percentage of TP orientation in open curves through a lesser need to perform anticipatory glances since the view up the road is unobstructed. Kandil et al. (2010) also consider looking “up the road” as an alternative to looking at the TP. Neither study, however, presented quantitative data on this anticipatory behavior. For further discussion and a more quantitative analysis of these look-ahead fixations, see Lehtonen et al. (2013, 2014).

6Optokinetic nystagmus (OKN) is generally considered to be a compensating mechanism that reduces slip in the retinal image during global flow in the visual scene, caused by self-motion through a textured environment (e.g., when looking out of the window of a train). During curvilinear locomotion in three dimensions, there is no globally coherent flow field and the slow phase must thus be tracking only local flow, or even a focal target. It is unknown whether the OKN slow phase is driven by an optokinetic reflex—typically regarded as an automatic process driven by retinal slip of the image—or a smooth pursuit that is considered a more complex tracking process involving top-down prediction (Kowler, 1989; Krauzlis, 2004). Both types of processes may also be involved; as in this case their effect on oculomotor behavior is synergistic, rather than a conflict between different control mechanisms.

7Incidentally, it appears that at least in some conditions OKN is also present in bicycling (Vansteenwijk, 2013). So far it has not been reported in locomotion on foot, where it appears that people tend to (a) fixate their locomotor target, (b) fixate the turn point of a curved FP (tracking a waypoint, but no periodicity), and (c) use a travel point moving along the FP some fixed distance or time-distance ahead on the path (no OKN) (see Imai, Moore, Raphan, & Cohen, 2001; Marigold & Patla, 2007; Bernardin et al., 2012). More data on walking and cycling on more realistic steering tasks (not constrained to 10–15 s, with the entire path visible before setting off, or constrained to make turns at visually designated points, or following a very narrow “lane”) are needed.

8Sometimes referred to as taking a “racing line,” but this is not really correct. The cutting behavior in normal driving seems more likely to involve turning into the curve very early (perhaps in order to reduce uncomfortable jerk in the onset of rotation). A true racing line, in contrast, involves a relatively late and a very fast initial
steering action. Cutting to the apex too early is one of the things one needs to unlearn in track driving.

References


Krauzlis, R. J. (2004). Recasting the smooth pursuit eye


