Transportation Research Part F 26 (2014) 258–272

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

Driving context and visual-manual phone tasks influence glance behavior in naturalistic driving

Emma Tivesten [a,b,*], Marco Dozza [a]

[a] *Chalmers University of Technology, Sweden*
[b] *Volvo Car Corporation, Sweden*

ARTICLE INFO

Article history:
Received 27 January 2014
Received in revised form 7 August 2014
Accepted 9 August 2014

Keywords:
Driver behavior
Eye movement
Gaze
Driving context
Phone tasks
Distraction

ABSTRACT

Naturalistic driving studies show that drivers engaged in complex visual-manual tasks face an increased risk of a crash or near-crash. Tasks that require many glances and a high proportion of long glances away from the road are of special concern for safety. Driving context (e.g. turning maneuvers, presence of lead or oncoming vehicles, vehicle speed) may also influence drivers' glance behavior during normal driving, since the drivers may have to estimate curvature and anticipate potential threats (e.g., lead vehicle braking). However, the effect of driving context on glance behavior during visual-manual tasks has not yet been thoroughly investigated in naturalistic driving. The extent to which drivers adapt their glance behavior to changes in the road environment during secondary tasks is likely to influence their ability to compensate for and respond to changes in the road environment. The present study investigated for the first time the effect of both driving context and visual-manual phone tasks (i.e., dialing, texting, reading) on drivers' glance behavior in naturalistic driving.

This study shows that drivers indeed spend more time looking at the road and have a lower proportion of long off-road glances in complex driving contexts such as when turning and when lead or oncoming vehicles are present, both in normal driving and while engaged in a visual-manual phone task. In particular, these findings are more pronounced during turning maneuvers and in the presence of oncoming vehicles than in the presence of lead vehicles. Interestingly, driving speed influenced off-road glance durations during the phone tasks, but not during normal driving.

The results from this study highlight the need to take driving context into account when evaluating the influence of different secondary tasks, in-vehicle user interfaces and glance metrics on driving safety, including the risk of crash involvement. The finding that glance behavior is context-dependent in a naturalistic setting has further implications for distraction detection algorithms, driver support systems, and driver training. Finally, driving contexts should be matched when comparing glance behavior, while driving with and without secondary tasks.

1. Introduction

Driving is a complex task that relies heavily on visual information (Sivak, 1998). Looking away from the road at the wrong time can therefore have severe consequences. Driver distraction has been acknowledged as one of the leading causes of

---

* Corresponding author at: Volvo Car Corporation, Sweden.

http://dx.doi.org/10.1016/j.trf.2014.08.004
1369-8478/© 2014 The Authors. Published by Elsevier Ltd.

crashes (Craft & Preslopsky, 2009; Dingus et al., 2006; Stutts, Reinfurt, Staplin, & Rodgman, 2001; Wang, Knipling, & Goodman, 1996). The present study examines how different driving contexts and visual-manual phone tasks influence the drivers' glance behavior in naturalistic driving.

1.1. Driver visual behavior in different driving contexts

Drivers constantly direct their gaze to different areas to obtain detailed information from the traffic environment. When and where drivers look is driven by expectations and experience with similar situations (top-down mechanism), as well as stimulus properties of an object such as movement, luminance, size, color and contrast (bottom-up mechanism) (Engström, Victor, & Markkula, 2013; Räsänen & Summala, 2000; Summala & Räsänen, 2000). In dynamic everyday tasks such making a sandwich or driving, fixations precede actions (Land, 2006; Land & Hayhoe, 2001; Land & Lee, 1994). For instance, drivers typically make look-ahead fixations to curves 1–2 s before entering the curve (Land & Lee, 1994).

Experienced drivers regularly fixate the road far ahead to control direction (Land, 2006), using their peripheral vision to control the lateral position of the vehicle (Land & Horwood, 1995; Mourant & Rockwell, 1970; Summala, Nieminen, & Punto, 1996). In on-road studies, drivers regularly fixate different areas of a curve to estimate curvature, control steering, and anticipate potential threats such as oncoming vehicles (Land & Lee, 1994; Lappi, Lehtonen, Pekkanen, & Itkonen, 2013; Lehtonen, Lappi, & Summala, 2012). Using the voluntary occlusion technique in a driving simulator, Tsimhoni and Green (1999) found that visual demand (as a proportion of viewing time) increases linearly as a function of increasing curvature (i.e. 1/radius).

Attending to other road users is also a crucial task for driving safely, in order to be able to respond to a lead vehicle braking or a pedestrian crossing, for instance. On-road studies have demonstrated that drivers look at the lead vehicle most of the time in car-following (Olson, Battle, & Aoki, 1989), and that drivers have long fixations to oncoming vehicles on rural roads (Serafin, 1994). Tijerina, Barickman, and Mazzae's (2004) on-road study of car-following found that drivers' glance behavior was not sensitive to speed or range, but the drivers tended to look away only when the optical expansion rate (and range rate) was very close to zero. Tijerina et al. (2004) compared their results to an earlier study by Rockwell (1988), and suggested that drivers' glance behavior is more cautious when a lead vehicle is present. Salvucci (2006) defined two constraints in a model of longitudinal control in car-following: keeping steady time headway and approximating desired time headway to the lead vehicle.

1.2. Driver visual behavior during visual-manual secondary tasks

Drivers use a time-sharing strategy when engaged in a visual-manual task where the gaze is constantly shifted between the secondary task and the driving scene for short intervals of time. Wierwille (1993a, 1993b) describes a model of visual time-sharing that suggests drivers perceive off-road glances that are shorter than one second as unproblematic, while they avoid off-road glances longer than 1.5 s. Several researchers have proposed that the driver gradually becomes less aware of the traffic environment over time as the secondary task continues (Lansdown, 2001; Zwahlen, Adams, & De Bald, 1988). This decreased awareness will only be partially improved during the short on-road glances that are common during visual time-sharing. Following this line of reasoning, long tasks would be more detrimental to awareness of the traffic environment than short tasks, even if the information can be chunked into a series of short off-road glances. Previous studies have found that different in-vehicle tasks vary somewhat for mean off-road glance duration, and considerably for the number of glances away from the road and the total glance time required to complete the task (Dingus, Hulse, Antin, & Wierwille, 1989; Wierwille, 1993b).

Both driving simulator studies and on-road experiments have shown that driving context also influences drivers' glance behavior during visual-manual tasks. Driving in a curve during a visual-manual task is associated with shorter off-road glances and a higher percentage of time looking at the road compared to driving on a straight road (Tsimhoni & Green, 2001; Victor, Harbluk, & Engström, 2005). Wierwille (1993a, 1993b) has shown that drivers' glance behavior is sensitive to traffic density and crosswind disturbance in experiments, but it is not understood how the presence of other vehicles influences drivers' glance behavior during visual-manual tasks in naturalistic driving. An increased understanding of this topic would improve the ability to detect and evaluate driver distraction in various driving contexts in real world driving, and more generally would also inform countermeasure development.

1.3. The safety relevance of different glance metrics

Mobile phone use is the most debated and studied form of driver distraction. Although talking on the phone seems to have no effect on crash risk, visual-manual phone tasks, such as dialing or texting, significantly increase crash and near-crash risk (Fitch et al., 2013; Hickman & Hanowski, 2012; Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006; Klauer et al., 2014; Olson, Hanowski, Hickman, & Bocanegra, 2009). In general, naturalistic driving studies have found that both complex visual-manual tasks and looking away from the road for more than two seconds (from one or more glances) in a six-second window increase crash and near-crash risk (Klauer et al., 2006; Olson et al., 2009). Liang, Lee, and Yekhshatyan (2012) further analyzed the 100-car data in order to evaluate how well different algorithms could predict crash risk. They found that single off-road glance duration is a more powerful predictor of crash risk than glance history or glance location. Total glance time is commonly used to measure total visual demand from secondary tasks (Dingus et al., 1989; Lansdown, 2001), but it is currently unknown to what extent it is related to crash risk.

Several design guidelines, test methods, and glance metrics have been proposed to evaluate the potentially distracting effect of in-vehicle user interfaces or secondary tasks (Green, 2008; NHTSA., 2013). In March 2013, NHTSA released their design guideline for in-vehicle user interfaces (NHTSA, 2013). The test method involves subjects driving on a 4-lane undivided road with a 50 mph speed limit in a car-following scenario on a straight road segment. Evaluation metrics include total glance time, mean glance duration, and the percentage of off-road glances that exceed 2 s.

In summary, the current literature indicates that long off-road glances as well as the percentage of time looking at the road are both relevant for traffic safety. Drivers need to spend more time looking at the road when driving in curves and when other vehicles are present in order to control steering and anticipate potential threats (e.g., lead vehicle braking). The understanding of how different driving contexts influence drivers' glance behavior in naturalistic driving is, however, limited. Drivers' ability to adapt their glance behavior to changes in the road environment during visual-manual tasks may also influence their ability to respond to potential hazards. Context-dependent glance behavior may therefore be relevant when estimating crash risk, developing driver support systems, and evaluating in-vehicle user interfaces.

1.4. Objective

The objective of this study is to investigate how different driving contexts and visual-manual phone tasks influence drivers' glance behavior in naturalistic driving. Specifically, the following research questions were formulated: How is glance behavior influenced by: (1) phone task engagement? (2) different phone tasks (i.e., dialing, texting or reading)?, and (3) different driving contexts (i.e. turning maneuvers, the presence of lead or oncoming vehicles, car speed and weather conditions)? The influence of context is considered both in normal driving and while the driver is engaged in a visual-manual phone task.

2. Materials and methods

2.1. The Swedish EuroFOT database for passenger cars

This study analyzed data collected in the EuroFOT project, which included naturalistic data from 100 Volvo cars driven in real traffic for one year. Driving data were recorded continuously at 10 Hz from engine ignition to engine stop, and included signals from cameras (forward road view, rearward road view, driver view), on-board sensors, and the CAN-bus. In total, approximately 1.0 million km were recorded and stored in the database, comprising 198 drivers (M = 45.3 years, SD = 10.8 years, 57% male, 43% female). See Sanchez et al. (2012) on Swedish VMC and passenger cars for more details on the collected data, including driver demographics.

2.2. Selection of whole trips and preliminary coding

A 5-week period during late spring 2010 was targeted for analysis, consisting of approximately 6000 trips. This period was selected because no driver support systems were activated and it did not include the summer holiday period. Three analysts viewed video-recordings of whole trips to identify any phone tasks (i.e., dialing, texting or reading). The trips were randomized from the database, and the coding work continued until the end of a pre-defined work period. Trips which were extremely short, were missing forward or driver video, or did not entail driving (e.g., car wash, ferry) were excluded. Only phone tasks that lasted for at least three seconds were included. In total, 366 trips were viewed and 49 of these trips contained at least one dialing, texting and/or reading task. The analysts made a preliminary coding of the variable *Phone-related sequence* (further described in Section 2.3) whenever they identified sequences containing dialing, texting or reading on a phone.

2.3. Selection, classification and coding of phone-related sequences

For each visual-manual phone task, a phone-related sequence was selected. Each sequence included 30 s of driving before the phone task, one phone task, and (in some cases) intervals during which the task was momentarily suspended. The variable *Phone-related sequence* was coded as a time series, and it contained the mutually exclusive categories: *Baseline, Dialing, Texting, Reading, Texting or Reading*, and *Suspend*. Most of the driving time was not captured by these categories and was consequently not included in further coding or analysis.

The 30 s driving period immediately before each phone task was categorized as *Baseline*. Searching for and picking up the phone was included in the baseline, since it was considered a different type of secondary task (i.e., reaching for/picking up object) rather than dialing, texting or reading. The baseline period was shorter than 30 s if the driver initiated the phone task shortly after another phone activity, or shortly after the recording of the trip started. Consequently, the baseline sequences consisted mostly of attentive driving, and occasionally included various secondary tasks (excluding the specifically coded visual-manual phone tasks or simply talking on the phone). The baseline sequences were considered representative of normal driving without any phone tasks, and were matched to the phone tasks which immediately followed, with respect to driving context (e.g., road type, lead vehicles).

A phone task involved dialing a phone number, writing and sending a text message or reading something on the phone. The phone tasks were categorized as either *Dialing*, *Texting*, *Reading*, or *Texting or Reading*. The distinction between phone tasks was based on the interaction with the phone buttons and the activities that followed after the phone interaction (e.g., holding phone to ear, waiting, talking, putting the phone away). The *Texting or Reading* category was used only when a task could not be defined as consisting solely of texting or of reading, either because (1) it was not possible to see if the driver interacted with the phone buttons while reading the phone screen, or (2) the driver was using the phone for reading and texting without a pause between the two tasks (which was then considered as one task). A phone task started at the beginning of the first off-road glance directly related to the task (ISO, 2002), and stopped at the end of the last off-road glance related to the task.

During a phone task, whenever the driver paused for at least three seconds or put the phone away, the task was categorized as *Suspend*. These periods of suspension were not included in the analysis, since the drivers clearly put the phone task on hold to attend to the driving task more closely.

2.4. Detailed coding performed for phone-related sequences

One person performed all the detailed video coding for the phone-related sequences using an updated version of FOTware (Dozza, Moeschlin, & Léon-Cano, 2010), a tool for the analysis of naturalistic driving data. Each sequence was coded according to these variables: *Eyes on road*, *Turning*, *Road type*, *Lead vehicle present*, *Oncoming vehicle present*, and *Weather conditions*. These variables, further described below, contain mutually exclusive categories. The sequences were coded as time series with 10 Hz resolution. A total of about 1 h and 30 min of driving was coded.

Eyes on road was defined as a binary variable categorized as *On road* or *Off road*. It was not coded if a clear view of the driver's eyes was missing (e.g., dark sunglasses). Glances were considered *On road* if the driver was looking at the road ahead, at intersecting roads, or at other road users relevant for the traffic situation, including looking in rear or side view mirrors. The rationale for coding glances at the mirrors as *On road* was that they were highly related to the driving task, since the driver was watching the road and checking for other vehicles when preparing to change lanes or overtake another vehicle. This behavior is also associated with a clearly reduced crash risk (Klauer et al., 2006). However, when a glance at a mirror was not related to the driving task (e.g., applying make-up), it was considered to be eyes *Off road*. The start and end of each glance was defined according to ISO (2002).

The variable *Turning* was defined as a binary variable categorized as *Yes* or *No* to indicate whether the driver was turning or not. The turning category included turning at intersections or in curves with a radius less than 1000 m. Four data sources were combined to estimate curve radius as well as the points in time when entering and exiting the curve. These sources were forward video, yaw rate (estimating radius according to Eq. (1)), maps[1] and GPS.

$$R_{CURVE} = V_{CAR}/(\omega * \pi/180) \tag{1}$$

In Eq. (1), ω is the yaw rate in degrees/s, V_{CAR} is the car velocity in m/s and R_{CURVE} is the estimated curve radius in meters.

The variable *Road type* was categorized as *Motorway/Highway*, *Rural*, *City* and *Other*. Roads where the posted speed limit was above 50 km/h were categorized as *Motorway/Highway* if there was a median barrier present, and as *Rural* if there was no median barrier. Roads where the posted speed was 50 km/h or lower were categorized as *City*. Parking areas, driveways and constructions zones were categorized as *Other,* independent of posted speed limit. This variable was specifically coded to identify rural roads without a median barrier when analyzing the presence of oncoming vehicles.

Lead vehicle present was coded by using a combination of the forward camera view and the distance to the vehicle ahead from the forward radar. The variable was coded as *Yes* if a vehicle ahead was within 150 m and traveling in the same lane as the subject vehicle, and otherwise coded as *No.*

Oncoming vehicle present was coded as *Yes* if a vehicle was visible in the front view camera and less than three seconds away from passing the subject car. Oncoming vehicles were not considered a relevant threat when a median barrier was present, so these were coded as *No*.

Weather conditions were coded as *Rain*, *Snow, Fog* or *Clear* using the forward view camera. The categories *Rain* and *Snow* were coded when there was precipitation, and *Fog* was coded anytime fog reduced visibility. All other cases were coded as *Clear*. Light conditions were not considered in this study, since the vast majority (more than 95%) of the phone tasks were performed in daylight.

2.5. Analysis

All parts of the phone-related sequences where the subject car was standing still (speed < 0.1 km/h) were filtered out after the coding and prior to the analysis, since the drivers' glance behavior was expected to be different when the drivers performed a phone task while the car was at a standstill.

---

[1] http://maps.google.se, http://www.eniro.se.

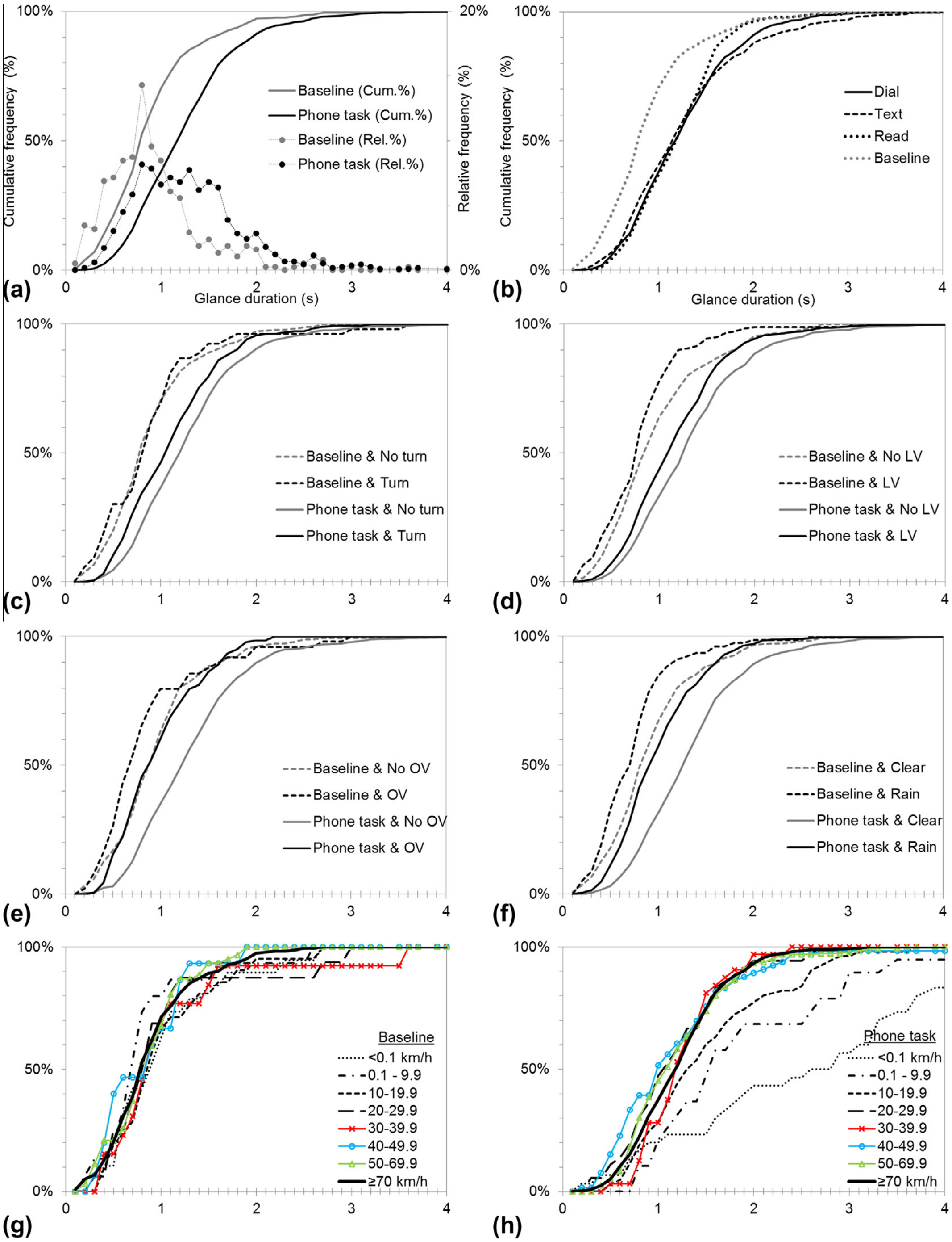

Fig. 1. Cumulative (a–h) and percent (a) frequency distributions of off-road glance duration in baseline driving and while drivers are engaged in a phone task: (a) all baseline driving vs. phone task operation, (b) different phone tasks, (c) turning or not turning, (d) lead vehicle present or not, (e) oncoming vehicle present or not, (f) weather conditions, (g) vehicle speed in baseline and (h) vehicle speed while drivers are engaged in a phone task. Note that standstill is included for comparison in figures (g) and (h), while standstill is not included in the analysis.

Different driving contexts (e.g., turning, lead vehicle presence) and phone task engagement were investigated by plotting the cumulative percent distributions of off-road glance durations. This provided a preliminary descriptive analysis about how different driving contexts influenced the shape of the glance distributions during phone task engagement.

Off-road glance frequency distributions for different glance durations and cumulative off-road glance durations were also plotted for the different phone tasks, so as not to overlook the influence of total glance time. The corresponding baseline data were plotted for comparison, where the baseline data were scaled to correspond to the mean duration of the phone tasks. In Fig. 2a, for instance, the glance frequencies and cumulative glance time from the baseline segments were scaled with a factor equal to the mean total task time for the dialing tasks divided by the mean duration of the baseline segments.

The following glance metrics were defined: Percentage of time that the driver spent looking at the road (%EONR), maximum off-road glance duration (MaxGD), percentage of off-road glances that were equal or longer than 2.0 s (%GD ⩾ 2s), the number of off-road glances or glance frequency (GF), and total glance time (TGT). TGT is the sum of all off-road glance durations during each phone task and is only considered for complete phone tasks. %EONR was computed for all segments (e.g., a curve segment in baseline) prepared for statistical comparisons, further described below. MaxGD, %GD ⩾ 2s and GF were computed for segments that included off-road glances. Here, GF is defined as the total number of glances within each analysis segment. Further, a segment was included in the analysis only if the duration was at least three seconds. Baseline segments were shorter than three seconds if they occurred at the very beginning of the trip recording, or the driver initiated a phone task less than three seconds after ending a previous phone call or visual-manual phone task.

Both the original metrics and their LOG10 transformations failed to meet the assumptions of normality and homogeneity of variance for all comparisons, described below. Non-parametric tests were therefore performed for all comparisons.

Different statistical tests were performed, depending on whether the context generally remained the same during a whole sequence (e.g. lead vehicle presence) or changed several times during a sequence (e.g. oncoming vehicle presence). Using the defined glance metrics, the following statistical comparisons were performed to verify the effects of phone tasks and driving contexts on glance behavior:

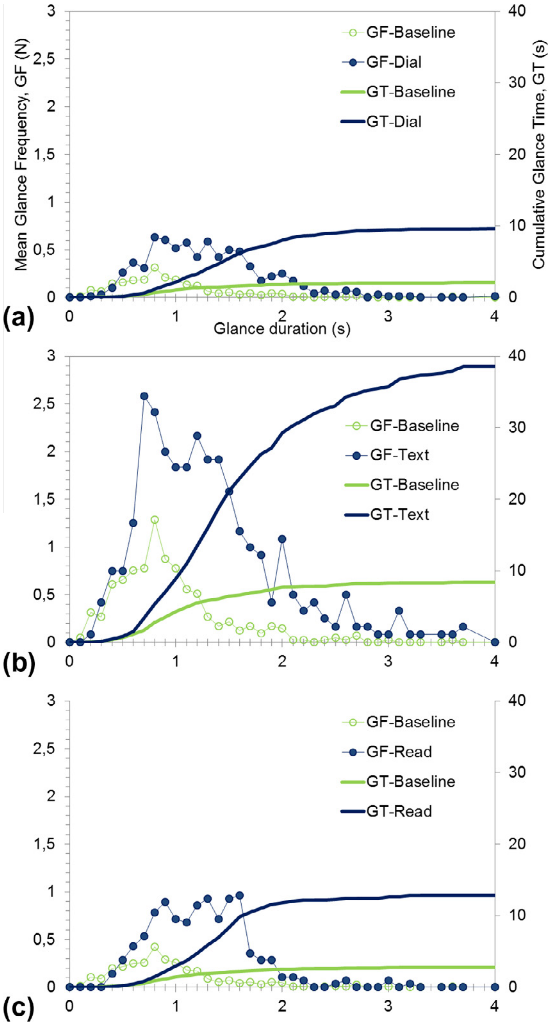

Fig. 2. Distribution of mean glance frequency (GF) on the primary y-axis, and the cumulative glance time (GT) on the secondary y-axis for different off-road glance durations on the x-axis. The blue bullets show mean glance frequency per task for all dialing (a), texting (b) and reading (c) tasks. The cumulative glance time is plotted as blue solid lines and corresponds to cumulative value of GF multiplied with glance duration. The baseline segments are scaled to correspond to the same mean task duration as the phone task in a–c respectively and are shown in light green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

- *Baseline* and phone task conditions (*Dialing, Reading, Texting,* and *Texting or Reading)* were compared within the same phone-related sequence using the Wilcoxon signed-rank test.
- The differences between *Dialing, Texting,* and *Reading* were investigated using the Kruskal–Wallis test. The Mann Whitney U post hoc test verified the individual differences between phone tasks whenever the Kruskal–Wallis test was statistically significant ($p < 0.05$).
- Turning at intersections or in curves was compared to not turning, in *Baseline* and in the phone task conditions. The Mann Whitney U test was used to compare different phone-related sequences.
- *Lead vehicle present* was compared to not present, in *Baseline* and in the phone task conditions using the Mann Whitney U test.
- *Oncoming vehicle present* was compared to not present on rural roads, in *Baseline* and in the phone task conditions using the Wilcoxon signed-rank test.
- *Clear* weather conditions were compared to *Rain*, in *Baseline* and in the phone task conditions using the Mann Whitney U test.

The calculated *p*-values were corrected for multiple tests using the Benjamin and Hochberg false discovery rate (FDR) (Benjamini & Hochberg, 1995) and were reported as p_c. In addition to the post hoc tests, a total of 26 other statistical tests, which addressed the different metrics and contexts, were included in the FDR corrections. The *p*-values for the post hoc tests were corrected with the FDR-factor for the Kruskal–Wallis test, and with Bonferroni corrections for the three types of phone tasks.

An additional analysis was performed to investigate the relation between vehicle speed and driver glance behavior. Spearman's correlation coefficient (r_s) was computed for mean speed and all glance metrics for both the baseline and phone tasks.

3. Results

In the 366 analyzed trips, visual-manual phone tasks were performed while driving in 49 of them, for a total of 109 tasks: 68 were coded as *Dialing*, 12 as *Texting*, 28 as *Reading*. One task was coded as *Texting or Reading* since it was impossible to see if the driver interacted with the phone buttons. All the phone tasks were performed using a hand-held phone, except for 3 dialing tasks in which the driver used the integrated phone keys in the center console. When considering total task duration, dialing sequences (Mdn = 10.4 s, M = 13.4 s, SD = 8.7 s) were similar to reading ones (Mdn = 11.2 s, M = 18.3 s, SD = 16.6 s), while texting sequences were longer (Mdn = 42.2 s, M = 55.2 s, SD = 39.1 s). Only 91 of the 109 baseline segments were available for analysis, because the others were shorter than three seconds.

3.1. Effect of phone task on glance behavior

Glance behavior was significantly different when the driver was also engaged in a phone task compared to baseline driving, for all evaluated glance metrics (Table 1). Off-road glances were clearly shifted toward longer glance durations while the driver was engaged in a phone task (Fig. 1a), indicating that drivers exhibited longer glances (significantly larger MaxGD, and %GD $\geqslant$ 2s; Table 1) and spent more time looking away from the road (significantly smaller %EONR; Table 1) when engaged in a phone task.

3.2. Effect of different phone tasks on glance behavior

Glance behavior varied for dialing, texting, and reading. The shape of the off-road glance distributions mainly differed among the phone tasks for glances longer than 1.5 s (Fig. 1b). In fact, there was a statistically significant difference between the three phone tasks when considering long glances (MaxGD and %GD $\geqslant$ 2s) (Table 1). Post-hoc tests revealed that texting had significantly more long glances (MaxGD, %GD $\geqslant$ 2s) than reading, and %GD $\geqslant$ 2s was also significantly higher for dialing than reading (Table 1). The phone tasks also exhibited a significant difference in the total number of glances (GF) and total glance time (TGT; Table 1). Texting had significantly higher GF and TGT than both dialing and reading (Table 1). The differences in TGT for the three phone tasks are shown in Fig. 2a–c, where the highest value of the cumulative glance time distributions corresponds to mean TGT. The baseline segments have shorter total glance times than the phone tasks, when scaled to the same mean duration as each phone task (Fig. 2a–c). In average, %EONR was largest for texting and smallest for dialing, although not statistically significant different across tasks (Table 1).

3.3. Effect of turning at intersections or in curves on glance behavior

In general, drivers spent more time looking at the road while turning compared to not turning, both at baseline and while they were engaged in a phone task. In fact, %EONR was significantly larger when turning than when not turning in baseline driving (Table 2). The off-road glance distributions are shifted toward shorter glances when turning, although the difference is less pronounced in baseline driving than during the phone tasks (Fig. 1c). While the driver was engaged

Table 1
Glance metrics for the three phone tasks.

	N_{EONR}/N_{GD}	%EONR Mdn (M, SD) Test statistic	MaxGD Mdn (M, SD) Test statistic	%GD ⩾ 2s Mdn (M, SD) Test statistic	GF Mdn (M, SD) Test statistic	TGT Mdn (M, SD) Test statistic
Baseline (BL)	91/85	88.0 (84.6, 12.4)%	1.10 (1.28, 0.61) s	0.0 (3.4, 9.7)%	4.0 (4.4, 2.8)	N/A
Phone task (PT)	109	26.2 (27.4, 10.3)%	2.00 (2.12, 0.68) s	3.4 (12.8, 18.6)%	7.0 (10.7, 12.0)	13.6 (8.5, 14.6) s
BL vs. PT [‡]	89/83 pairs	$z = -8.19$, $r = -0.61$ **p_c < .001**	$z = -6.92$, $r = -0.54$ **p_c < .001**	$z = -4.07$, $r = -0.32$ **p_c < .001**	N/A	N/A
Dial	68	24.8 (25.9, 10.0)%	2.00 (2.10, 0.66) s	8.7 (15.4, 19.9)%	6.0 (7.5, 5.2)	8.2 (9.7, 6.0) s
Text	12	34.5 (30.6, 9.1)%	2.50 (2.73, 0.85) s	9.3 (17.6, 21.2)%	27.5 (29.7, 23.3)	30.7 (38.6, 27.5) s
Read	28	30.0 (29.8, 11.4)%	1.80 (1.89, 0.50) s	0.0 (4.8, 10.7)%	5.5 (10.3, 10.3)	6.9 (12.7, 11.7) s
Comparisons between phone tasks [§]	68 12 28	$H(2) = 4.38$ $p_c = 0.161$	$H(2) = 11.47$, **p_c = .010** Dial-text (n.s.) Dial-read (n.s.) **Text-read**[*]	$H(2) = 10.35$, **p_c = .016** Dial-text (n.s.) **Dial-read**[*] **Text-read**[*]	$H(2) = 21.95$, **p_c < .001** **Dial-text**[***] Dial-read (n.s.) **Text-read**[**]	$H(2) = 21.48$, **p_c < .001** **Dial-text**[***] Dial-read (n.s.) **Text-read**[**]

BL = Baseline, PT = phone task, N_{EONR} = Total number of segments, N_{GD} = Number of segments with off-road glances %EONR = % of time with eyes on road, MaxGD = Maximum off-road glance duration, %GD ⩾ 2s = % of glances ⩾ 2s, GF = Glance frequency, TGT = total glance time.
Bold font indicates statistically significant differences.

[‡] Wilcoxon signed-rank test.

[§] Kruskal–Wallis test and Mann Whitney U post hoc test: n.s. = not significant.

[*] $p_c < .05$.

[**] $p_c < .01$.

[***] $p_c < .001$.

Table 2
Glance metrics for the driving contexts: Presence of turning maneuvers, lead vehicle and oncoming vehicles.

	N_{EONR}/N_{GD}	%EONR Mdn (M, SD) Test statistic	MaxGD Mdn (M, SD) Test statistic	%GD ⩾ 2s Mdn (M, SD) Test statistic	GF Mdn (M, SD) Test statistic
BL & Turn	39/22	93.6 (88.6, 13.5)%	1.00 (1.14, 0.64) s	0.0 (3.0, 14.2)%	2.0 (2.2, 1.3)
BL & No Turn	89/82	87.0 (83.6, 13.0)%	1.10 (1.23, 0.57) s	0.0 (3.3, 9.6)%	3.0 (3.9, 2.6)
PT & Turn	31	30.5 (31.5, 11.5)%	1.70 (1.78, 0.55) s	0.0 (8.7, 16.4)%	4.0 (5.8, 5.0)
PT & No Turn	99	26.3 (27.3, 10.9)%	2.00 (2.11, 0.70) s	5.3 (13.7, 19.7)%	6.0 (9.7, 9.8)
BL & Turn vs. BL & No Turn[†]	39/22 52/48	$U = 724.0$, $z = -2.34$ **p_c = .036**, $r = -0.24$	$U = 464.5$, $z = -0.81$ $p_c = .480$, $r = -0.10$	$U = 457.5$, $z = -1.47$ $p_c = .210$, $r = -.18$	N/A
PT & Turn vs. PT & No Turn[†]	31 77	$U = 894.0$, $z = -2.03$ $p_c = .072$, $r = -.20$	$U = 828.0$, $z = -2.49$ **p_c = .029**, $r = -.24$	$U = 928.0$, $z = -1.97$ $p_c = .079$, $r = -.19$	N/A
BL & LV	47/42	89.0 (86.6, 10.8)%	1.05 (1.17, 0.52) s	0.0 (0.8, 3.8)%	3.0 (4.3, 3.0)
BL & No LV	54/49	87.0 (83.3, 13.5)%	1.20 (1.33, 0.66) s	0.0 (5.2, 12.0)%	4.0 (4.0, 2.6)
PT & LV	48	24.6 (26.5, 9.5)%	1.90 (2.18, 0.70) s	0.0 (10.3, 15.6)%	7.0 (12.0, 15.3)
PT & No LV	64	28.3 (28.2, 11.2)%	2.00 (2.03, 0.67) s	4.5 (14.7, 21.1)%	6.0 (9.1, 8.2)
BL & LV vs. BL & No LV[†]	47/42 44/41	$U = 846.5$, $z = -1.49$, $p_c = .188$, $r = -0.16$	$U = 685.0$, $z = -1.61$, $p_c = .165$, $r = -0.18$	$U = 706.0$, $z = -0.26$ **p_c = .022**, $r = -0.26$	N/A
PT & LV vs. PT & No LV[†]	48 61	$U = 1380.5$, $z = -.51$ $p_c = .613$, $r = -.05$	$U = 1303.5$, $z = -.98$ $p_c = .388$, $r = -.09$	$U = 1345.0$, $z = -.78$ $p_c = .456$, $r = -.07$	N/A
BL & OV	36/23	93.5 (86.9, 14.7)%	0.90 (1.11, 0.70) s	0.0 (11.6, 31.2)%	2.0 (2.0, 1.7)
BL & No OV	47/43	84.2 (80.5, 14.4)%	1.20 (1.38, 0.64) s	0.0 (4.4, 11.5)%	4.0 (4.2, 2.6)
PT & OV	22	35.0 (37.7, 13.1)%	1.70 (1.67, 0.35) s	0.0 (9.1, 24.0)%	4.0 (7.7, 10.6)
PT & No OV	55	21.4 (25.0, 10.6)%	2.00 (2.21, 0.78) s	7.7 (15.8, 21.0)%	6.0 (9.1, 9.4)
BL & OV vs. BL & No OV[‡]	34/19 Pairs	$z = -2.40$, **p_c = .031** $r = -.29$	$z = -1.49$, $p_c = .184$ $r = -.24$	$z = -0.94$, $p_c = .474$ $r = -.15$	N/A
PT & OV vs. PT & No OV[‡]	21 Pairs	$z = -3.01$, **p_c = .006** $r = -.46$	$z = -3.53$, **p_c < .001** $r = -.54$	$z = -2.42$, **p_c = .029** $r = -.37$	N/A

PT = phone task, LV = Lead vehicle, OV = Oncoming vehicle.
N_{EONR} = Total number of segments, N_{GD} = Number of segments with off-road glances.
%EONR = % of time with eyes on road, MaxGD = Maximum glance duration, %GD ⩾ 2s = % of glances ⩾ 2s, GF = Glance frequency.
Bold font indicates statistically significant differences.

[†] Mann Whitney U test.

[‡] Wilcoxon signed-rank test. BL = Baseline.

in a phone task, MaxGD was significantly shorter when turning than when not turning (Table 2). It is worth noting that sequences with off-road glance durations longer than 2 s during turns were quite rare in baseline driving, while more common during the phone tasks.

3.4. Effect of lead vehicle presence

The off-road glance frequency distribution was shifted toward shorter glance durations when a lead vehicle was present, both at baseline and while the driver was engaged in a phone task (Fig. 1d). During baseline driving, %GD ⩾ 2s was significantly lower when a lead vehicle was present (Table 2). On the other hand, lead vehicle presence did not yield statistically significant differences for any of the other glance metrics either in baseline driving or while the driver was engaged in a phone task (Table 2).

3.5. Effect of oncoming vehicle on rural roads

On rural roads, drivers generally spent more time looking at the road if an oncoming vehicle was present. In fact, %EONR was significantly larger when an oncoming vehicle was present both at baseline and while the driver was engaged in a phone task (Table 2). The glance frequency distribution in Fig. 1e is clearly shifted toward shorter durations when an oncoming vehicle is present, and this is more pronounced while the driver is engaged in a phone task than for baseline driving. In fact, during a phone task, the presence of an oncoming vehicle made a significant difference in MaxGD and %GD ⩾ 2s (Table 2). The glance frequency distribution during a phone task when oncoming vehicles are present is quite similar to the glance distribution at baseline when no oncoming vehicles are present, as shown in Fig. 1e.

3.6. Effect of weather conditions on glance behavior

More than 80% of the analyzed data consisted of trips made in clear weather conditions, and it was raining in the remaining cases. Drivers exhibited shorter glance durations while driving in the rain compared to clear weather conditions (Fig. 1f), both during baseline driving and while they were engaged in a phone task. There were, however, not enough data to show statistically significant differences for any other glance metrics for different weather conditions.

3.7. Effect of subject car speed on glance behavior

There was no clear influence of speed on the drivers' visual behavior during baseline driving, as shown in Fig. 1g, where the cumulative off-road glance frequency distributions for different speed intervals overlap. The glance metrics were not significantly correlated with speed during baseline driving (Table 3). On the other hand, the drivers' glance behavior was sensitive to speed when they were engaged in a phone task, particularly in speeds below 20 km/h, while the glance distributions in speed intervals above 20 km/h are more uniform (Fig. 1h). In fact, mean speed was significantly correlated with MaxGD and %GD ⩾ 2s when drivers were engaged in a phone task (Table 3).

4. Discussion

This study demonstrated and quantified the influence of driving context (e.g., turning maneuvers, presence of other vehicles, vehicle speed) and visual-manual phone tasks (i.e. dialing, reading or texting) on drivers' glance behavior in naturalistic driving.

4.1. Effect of driving context

Drivers adapted their glance behavior to the driving context, both during baseline driving and while engaged in a visual-manual phone task. In general, drivers spent more time looking at the road and had shorter off-road glances when turning,

Table 3
Glance metrics for different subject car speed intervals.

	N_{EONR}/N_{GD}	%EONR Mdn (M, SD) Test statistic	MaxGD Mdn (M, SD) Test statistic	%GD ⩾ 2s Mdn (M, SD) Test statistic	GF Mdn (M, SD) Test statistic
BL & 0.1–19.9 km/h	9/7	91.7 (85.6, 18.2)%	0.80 (1.23, 0.71) s	0.0 (3.6, 9.4)%	3.0 (4.4, 2.5)
BL & 20–39.9 km/h	10/9	91.7 (88.7, 9.0)%	1.10 (1.24, 0.59) s	0.0 (2.8, 8.3)%	2.0 (2.9, 1.5)
BL & 40–69.9 km/h	18/18	88.9 (83.8, 12.6)%	1.15 (1.37, 0.77) s	0.0 (2.6, 7.7)%	4.5 (4.9, 3.2)
BL & ⩾ 70 km/h	54/51	86.7 (83.9, 12.0)%	1.10 (1.27, 0.55) s	0.0 (3.8, 10.7)%	4.0 (4.5, 2.8)
Correlation with mean speed	91/85	$r_s = -.060$, $p = .569$	$r_s = -.059$, $p = .590$	$r_s = -.095$, $p = .389$	N/A
PT & 0.1–19.9 km/h	9	19.6 (24.4, 9.2)%	2.60 (2.46, 0.91) s	16.7 (18.0, 20.4)%	6.0 (9.0, 10.3)
PT & 20–39.9 km/h	11	24.3 (26.5, 7.7)%	2.00 (2.24, 0.62) s	16.7 (14.8, 12.6)%	5.0 (13.9, 27.0)
PT & 40–69.9 km/h	18	27.5 (27.3, 12.2)%	2.05 (2.32, 0.96) s	2.8 (13.3, 19.6)%	5.5 (11.6, 12.3)
PT & ⩾ 70 km/h	71	28.1 (28.0, 10.4)%	1.90 (2.00, 0.55) s	0.0 (11.7, 19.1)%	7.0 (10.1, 8.2)
Correlation with mean speed	109	$r_s = -.058$, $p = .548$	$r_s = -.287$, $\mathbf{p = .002}$	$r_s = -.192$, $\mathbf{p = .045}$	$r_s = -.012$, $p = .904$

BL = Baseline, PT = Phone Task, N_{EONR} = Total number of segments, N_{GD} = Number of segments with off-road glances.
%EONR = % of time with eyes on road, MaxGD = Maximum glance duration, %GD ⩾ 2s = % of glances ⩾ 2 s, GF = Glance frequency, TGT = total glance time.
Bold font indicates statistically significant differences.

when a lead vehicle was present, or when oncoming vehicles were present. These results are consistent with previous on-road and simulator experiments, which reported that drivers adapted their glance behavior to driving task demand (Wierwille, 1993a, 1993b). As a consequence, the same secondary task may need to satisfy more stringent requirements to be safely performed in a curve rather than on a straight road.

While turning, %EONR was significantly higher during baseline driving, and MaxGD significantly lower during the phone tasks. This behavior is probably driven by the higher visual demand while turning, and suggests that the driver anticipates potential hazards such as oncoming vehicles (Land & Lee, 1994; Lehtonen et al., 2012). The results are in line with previous simulator studies demonstrating that drivers' glance behavior is influenced by road curvature during visual-manual tasks (Tsimhoni & Green, 2001; Victor et al., 2005).

Drivers have slightly shorter glances and focus more on the road when a lead vehicle is present, both in baseline driving and during visual-manual phone tasks. Lead vehicle presence had less influence on drivers' glance behavior than turning maneuvers and oncoming vehicles. A possible explanation may be that drivers are more attentive to situations that constantly change (i.e. oncoming vehicles) than to steady-state driving situations, such as car-following when traffic flows freely for a long time. This interpretation is supported by on-road experiments and car-following models suggesting that driving demand and glance behavior are mainly influenced by changes in headway to the lead vehicle (Salvucci, 2006; Tijerina et al., 2004). Hence, maintaining constant headway in a monotonous situation at a fairly constant speed may give drivers a false sense of safety compared to driving situations that constantly change. As a result, drivers may be more susceptible to distraction and inattention. Thus, while engaged in visual-manual tasks in car-following, the driver may potentially be less prepared for sudden changes, such as the lead vehicle braking.

The results in this study clearly show that drivers look more at the road ahead when an oncoming vehicle is present; this behavior was evident both in baseline driving and during visual-manual phone tasks. Furthermore, drivers have shorter off-road glances during the phone tasks, as demonstrated by MaxGD and %GD $\geqslant$ 2s being significantly lower when an oncoming vehicle is present. The effect of oncoming vehicles on drivers' glance behavior is a novel finding which has not been demonstrated in previous studies. On rural roads (without a median barrier), drivers' glance behavior is more sensitive to the presence of oncoming vehicles than to the presence of a lead vehicle. This is supported by the more general finding by Wierwille (1993a, 1993b) that traffic density influences glance behavior. The effects of lead and oncoming vehicle presence on glance behavior may have different explanations. First, an oncoming vehicle may be perceived as a more severe threat than a lead vehicle. Indeed, oncoming vehicles have a large relative velocity and collisions with oncoming vehicles are associated with a high crash severity (Björnstig, Björnstig, & Eriksson, 2008; Swedish Official Statistics., 2013). Second, oncoming vehicles show as a rapid motion in the drivers' visual field (especially if the traffic density is low), which may direct the driver's attention to the road. In fact, the driver's peripheral vision is sensitive to movement (Baker & Braddick, 1985) and an oncoming vehicle might trigger the driver's attention by a bottom-up mechanism (Engström et al., 2013). Third, it is possible that longer fixation times are required for oncoming vehicles than lead vehicles in order to estimate relative speed and direction. This possibility is supported by the results of a previous on-road experiment (Serafin, 1994).

Drivers' glance behavior may be sensitive to weather conditions. Drivers exhibited shorter off-road glances in rainy weather than in clear weather. However, this result was not statistically significant and more data is needed to support this finding. Light condition may also influence off-road glance durations. However, the data set mainly comprised trips conducted in daylight, so it was not possible to assess the effect of light conditions on glance behavior. A recent driving simulator study showed that driving in rainy weather or dark conditions produces more and longer fixations on the road than driving in daylight during clear weather conditions (Konstantopoulos, Chapman, & Crundall, 2010). However, it is worth noting that the data set in the present study was collected during late spring in the Gothenburg region. There would be a much wider variation in weather (e.g., rain, snow, clear) and light conditions if data had covered a whole year.

Several previous studies on drivers' eye glance behavior performed in driving simulators or on-road studies focused on highway driving. Tijerina et al. (2004) specifically studied car-following, while other studies did not specify the presence of other vehicles (Rockwell, 1988; Victor et al., 2005). Given that driving context varied in different studies, it is difficult to compare their results for glance behavior. Our results show examples of context variables that should be controlled for in future research to facilitate comparisons across studies.

4.2. Effect of driving speed on glance behavior

Another interesting and novel finding from this study is that speed does not influence the drivers' glance behavior during baseline driving. However, drivers engaged in a phone task have longer off-road glances at lower speeds. This is especially evident for speeds below 20 km/h, while speeds above 20 km/h have more similar distributions for off-road glance durations, independent of speed. In baseline driving, most tasks can be performed using mainly short off-road glances. It is possible that glance duration below a certain threshold (e.g. 1 s) could be perceived as unproblematic independent of speed, in accordance with Wierwille's model of time-sharing (1993a, 1993b). Also, the effect of speed on phone tasks is to be expected when considering the low driving demand for lane keeping and car heading at low speeds. On the other hand, speeds below 20 km/h are common in stop-and-go traffic, where the probability of a lead vehicle braking is quite high, although the crash severity is typically low in case of a crash. The results suggest that, at slow speeds, drivers prioritize the secondary task over monitoring potential hazards in the traffic environment. Future research, including larger datasets, should further investigate the influence of speed on drivers' glance behavior to establish if this finding holds for a larger population.

4.3. Effect of phone tasks on glance behavior

Not surprisingly, the drivers spent less time looking at the road and had longer off-road glances while engaged in a visual-manual phone tasks compared to baseline driving, corroborating results from previous studies (Liang & Lee, 2010; Victor et al., 2005; Zhang, Smith, & Witt, 2006).

Among the phone tasks, there is a clear difference in the prevalence of long glances (as demonstrated by MaxGD and %GD $\geqslant$ 2s) as well as the glance frequency and total glance time. Previous simulator and on-road studies show that glance frequency, glance duration, and total glance time vary between various in-vehicle tasks (Dingus et al., 1989; Lansdown, 2001; Wierwille, 1993b). The total glance time for texting in this study is much longer than for any of the in-vehicle tasks investigated (including text entry) in previous experiments (Dingus et al., 1989; Wierwille, 1993b). This difference may be due to the naturalistic setup of this study. The off-road glance distributions in the present study for dialing, texting and reading differ mainly for glances longer than 1.5 s, corresponding to approximately the 75% ile glance duration. This is in agreement with studies indicating that the proportion of glances exceeding a certain threshold (e.g., 1.6 or 2 s) is more sensitive to the complexity of the task than to mean glance duration (Horrey & Wickens, 2007; Victor et al., 2005). The threshold of 1.5 s may be a more useful alternative to the 2 s threshold which may also vary depending on the driving task demand, secondary task demand, and individual differences (Wierwille 1993a, 1993b).

4.4. Sensitivity of different glance metrics

A central finding was that GF and glance metrics capturing long glances (MaxGD, %GD $\geqslant$ 2s) are the best metrics for distinguishing between different phone tasks. The long glance metrics are the most sensitive to driving contexts during visual-manual phone tasks. On the other hand, %EONR was more effective than the long glance metrics at distinguishing between different driving contexts in baseline driving. Consequently, %EONR may serve as a basic metric of visual allocation in all driving contexts, with or without secondary task involvement. However, the metrics capturing long glances and glance frequency may be more informative during visual-manual tasks. In this respect, the results presented in this paper agree with previous research regarding the sensitivity of different glance metrics for different tasks and driving contexts performed in simulator or on-road experiments (Victor et al., 2005).

4.5. Safety implication of different glance metrics

Sending a text message entails a higher proportion of long glances and a longer total glance time than dialing or reading. Texting is also associated with a considerably higher risk of crash and near-crash involvement than other visual-manual tasks (Olson et al., 2009). Our results suggest that this increased risk may be explained by the high proportion of long glances and the long total glance time (TGT). A long TGT has two implications for traffic safety. First, since these tasks take a long time to complete, the total time (exposure) dedicated to a single task can be fairly long, even if drivers perform this type of task less often than others (Wierwille & Tijerina, 1998). Second, it is possible that a driver becomes gradually less aware of the traffic environment while engaged in visual-manual task (Lansdown, 2001; Zwahlen et al., 1988). On the other hand, a recent study by Liang et al. (2012) has demonstrated that single glance duration is the most effective metric in naturalistic driving data for predicting crash risk, and considering glance history as well does not greatly increase the effectiveness, thus reducing the importance of TGT as a predictor for risk. This result from Liang et al. (2012) is in agreement with the model proposed by Horrey and Wickens (2007), in which crash risk is described as a function of the percentage of glance durations exceeding 1.6 s and the context variables describing the characteristics of the traffic situation.

4.6. Safety evaluation of visual-manual secondary tasks

Glance behavior during the three phone tasks can be related to the acceptance criteria in NHTSA design guidelines (NHTSA, 2013). From the results in Table 2 it is clear that more than half of the visual-manual phone tasks cannot be completed within a TGT less than 12 s, and most tasks have glance durations longer than two seconds. However, the NHTSA guidelines do not apply specifically to hand-held phone tasks, and the tasks specified in the guidelines are evaluated in a driving simulator in a specific driving context rather than in a naturalistic setting.

The results from the present study highlight the need to take driving context into account when evaluating visual distraction due to different secondary tasks, including in-vehicle user interfaces, since the crash/near-crash risk of different phone tasks vary with driving task demand (Fitch, Hanowski, & Guo, 2014b). One way to control for context is to describe a specific driving scenario, as in the NHTSA design guidelines (NHTSA, 2013). However, when evaluating the distracting effect of secondary tasks, it might be necessary to account for different driving contexts, since what can be considered safe glance behavior depends on the driving task demand. Evaluating secondary tasks in several different driving contexts is potentially important, since drivers do perform secondary tasks in varying situations. Further research exploring how different glance metrics influence crash risk could also improve future design guidelines. This would in turn promote the development of user interfaces and driver support functions that perform well from a real-life safety perspective.

4.7. Implications for distraction detection algorithms and countermeasures

Several studies, such as Kircher, Ahlström, and Kircher (2009) and Victor (2010), have proposed distraction detection algorithms based on drivers' eye glance behavior. The present study shows that driving context variables such as curvature and the presence of other vehicles should be added to existing algorithms, to improve the detection and assessment of visual distraction. For instance, glances longer than 2 s may impose a much lower crash risk while driving on a motorway without a lead vehicle present, than while following a lead vehicle with short time headway in high-density traffic or when negotiating a sharp curve. Furthermore, the modest influence of driving speed on off-road glance duration may also have some implications for the development of future algorithms. Some of the existing algorithms only assess driver distraction above a certain speed, such as 40 or 50 km/h (Kircher et al., 2009; Victor, 2010). The results suggest that it might not be necessary to take a speed dependency into account, thus enabling algorithms to include a larger range of driving conditions.

Detecting driver distraction from glance behavior and driving context could significantly improve current driver support systems. For instance, a forward collision warning could be more sensitive (i.e. warn earlier) if the driver is engaged in a visual-manual task. A prerequisite for such a system would be a reliable measurement and assessment of both the driver's on/off-road glances and the road environment. Driver post-drive feedback, or real-time feedback about their glance behavior (Lee et al., 2013), is another type of driver support system that could benefit from the inclusion of the context variables in this study.

Legislation and driving education are other safety countermeasures which may take advantage of the results presented in this study. For instance, phone interaction could be prohibited on certain road areas (e.g., curved stretches of road or approaches to traffic lights in unexpected locations) where the driver would be most likely to face difficulties while multitasking. These data were collected in Sweden, where there is no legislation against mobile phone use while driving. It is possible that a ban (full or conditional) would affect mobile phone use while driving (McCartt, Hellinga, Strouse, & Farmer, 2010). In addition, driver education could include making new drivers aware of safe glance behavior in different driving contexts.

4.8. Strengths and limitations of naturalistic driving data

The main advantage of using naturalistic driving data for the analysis of driver behavior is that the collected data shows genuine driver behavior, without the influence of artificial environments, experimenters, or experimental protocols which are to be expected in simulators or test tracks. There are, however, limitations to the use of naturalistic data and the present study that are worth mentioning. Naturalistic data is often geographically and demographically biased. In this study the participant sample was biased, as the primary drivers were mainly males, 35–65 years of age. However, the fact that other members of the household also drove the cars made the driving data somewhat more representative of the driving population. In addition, all the data was collected in one Scandinavian city. The present study has taken phone task and driving context into account for a group of individuals that can be considered experienced drivers. Since there are many drivers in a dataset of limited size, the results should be interpreted as valid for experienced drivers in general. The data available for each driver is, however, too limited to study individual differences. Previous studies show that drivers' glance behavior differs for novice and experienced drivers (Wikman, Nieminen, & Summala, 1998), for drivers' age (Dukic, Hanson, & Falkmer, 2006), and that some individuals are consistently "long glancers" (Broström, Ljung Aust, Wahlberg, & Källgren, 2013). Even though video coding may currently be more reliable than eye tracker data in naturalistic driving studies (Ahlström, Victor, Wege, & Steinmetz, 2012), it has some limitations worth noting. First, the coding is based on video recordings at 10 Hz, providing a 0.1 s resolution which potentially biases short glances. Second, it is not possible to establish precisely when fixations on the phone end and a saccade back to the road begins. In some cases, part of the transition time from the phone to the road may be included in off-road glance duration. This means that the coded off-road glance duration in this study may be slightly longer than the formulated ISO standard definition of a glance (ISO, 2002). Third, all coding were performed by a single person, and inter-rater reliability is an important methodological issue. However, since Smith, Chang, Glassco, Foley, and Cohen (2005) found that raters agree most of the time when coding eyes on- or off-road from video-recorded data, it is reasonable to assume that other analysts trained in performing video-based coding would obtain similar results. It is, however, recommended to have several coders to check for inter-rater reliability, and to manage the work load of analyzing larger datasets. A larger dataset would increase the number of tasks available for analysis, and increase the confidence in the statistical tests. In particular, texting tasks were quite rare, only present on 12 occasions.

The selection of baseline driving segments has implications for the analysis. In this study, baseline segments consisted of 30 s prior to each phone task in order to achieve a baseline sample that was matched to the phone tasks (same drivers and similar driving context). The rationale for this choice was also that the baseline would represent normal driving including occasional secondary task engagement (except for phone tasks). It is, however, worth noting that reaching for the phone was generally included in the baseline segments. Therefore, reaching for an object is likely overrepresented in the baseline segments in comparison to driving in general. However, the reaching tasks did in most cases involve a short single glance just before picking up the phone. One alternative is to select baseline segments from driving that occurred before the drivers reached for the phone to perform a visual-manual phone task, as in Fitch et al. (2013), where reaching tasks generally had low TGT. This alternative approach would have increased %EOR and decreased GF in 30 s baseline segments slightly, presenting even larger differences between phone tasks and baseline segments. Excluding the reaching task from baseline is not obvious and require a more complex selection of the baseline segments since the driver sometimes performs the visual-manual tasks right after picking up the phone, and sometimes the driver picks up the phone and then wait for a good opportunity to perform the task.

4.9. Future research: the role of self-regulation on driver distraction

This paper demonstrates that driving context affects glance behavior both during normal driving and while drivers are engaged in dialing, texting, or reading on a mobile phone. A following question is whether glance behavior is affected similarly by context under the two driving conditions. In other words, are secondary tasks more distracting in more complex driving contexts? Although this question goes beyond the scope of this study, the results may provide some insights. In fact, a comparison of the dashed and solid curves in Fig. 1 according to their color demonstrates the effect of secondary task engagement in more or less complex driving contexts. For instance, in Fig. 1c comparing the gray dashed and solid lines shows the effect of a phone task while turning, and the difference between the corresponding black lines shows the effect of a phone task while driving straight. The dashed lines in Fig. 1c are overlapping whereas the solid gray line is shifted more to the right than the black line, indicating clearly that the effect of a secondary task on glance behavior is higher when not turning. The same visual exercise in Fig. 1d–f may indeed suggest to the reader that the less complex the context, the more the glance distribution shifts toward longer glances away from the road during the phone tasks. If statistically proved in future studies, this result can be linked to the results in Fig. 1g–h, which show that the lower the speed the longer the glances away from the road are during the phone tasks. In fact, a less dangerous situation (lower speed or less complex context) may result in the driver allowing for more distraction. In other words, drivers may self-regulate their behavior and be less prone to distraction in more demanding driving situations.

Future studies should also determine the extent to which drivers plan for secondary tasks (e.g. by choosing a specific driving context to initiate a secondary task), which factors promote or trigger the decision to engage in secondary tasks, or if drivers compensate for secondary tasks as indicated in a few naturalistic driving studies (Fitch, Grove, Hanowski, & Perez, 2014a; Fitch et al., 2014b; Funkhouser & Sayer, 2012). If drivers decide to engage in secondary tasks only in specific contexts, then this decision-making should be taken into account to determine the influence of secondary tasks on driving, including crash risk. Currently, normal driving is extracted from naturalistic data by randomly picking baseline epochs, with the possible exclusion of standstills (Dingus et al., 2006; Klauer et al., 2006; Olson et al., 2009). However, if secondary tasks do happen in specific contexts, a fair comparison would require the normal driving sequences to be matched to secondary task sequences according to context. The same principle holds for design guidelines, which should test a secondary task in the most probable scenario.

5. Conclusions

Drivers adapt their glance behavior to the driving context, both during baseline driving and while attending to visual-manual phone tasks. Drivers spend more time looking at the road, and have a lower proportion of long off-road glances when turning, and when other vehicles are present. Drivers glance behavior is more sensitive to turning maneuvers or the presence of oncoming vehicles than to the presence of a lead vehicle. Interestingly, driving speed did not influence glance behavior during baseline driving, although it had a clear influence during phone tasks. Of the three phone tasks, texting resulted in a higher proportion of long glances and a much longer total glance time than dialing or reading.

All the investigated glance metrics can distinguish between phone tasks and baseline driving. The glance metrics capturing long glances (MaxGD and %GD $\geqslant$ 2s) are the best at distinguishing between tasks and driving contexts during visual-manual phone tasks. %EONR, in contrast, is more effective for distinguishing between different driving contexts in baseline driving. These results prove that several glance metrics and context variables are necessary to reliably assess drivers' visual distraction.

Future studies should include driving context variables in order to facilitate comparison of glance behavior across studies. Existing distraction detection algorithms could benefit from including context variables, not only to improve the detection of driver distraction, but also to assess the level of visual distraction. Enhanced distraction detection algorithms could significantly improve driver support systems. For instance, the sensitivity of a forward collision warning can be adjusted depending on glance behavior in relation to the context.

Further research is needed to improve our understanding of how drivers adapt their glance behavior to different driving contexts and the extent to which this adaptation controls for or facilitates distraction. Such knowledge should be included in driver models to support future distraction detection algorithms, driver support functions, and to inform the design of guidelines to evaluate distraction from secondary tasks.

Acknowledgments

This research was financially supported by the Swedish Governmental Agency for Innovation Systems (VINNOVA) and Volvo Car Corporation (VCC). The project was carried out within SAFER – Vehicle and Traffic Safety Centre at Chalmers University of Technology, Gothenburg, Sweden. The authors thank Anna Heyden, Ronja Örtlund, Caroline Gustafsson and Jonas

Attertun for performing the initial analysis of whole trips, as well as Sergejs Dombrovskis and Jordanka Kovaceva for their support with the EuroFOT database. The authors express their appreciation to their colleagues at SAFER, VCC and Chalmers, especially Prof. Trent Victor, Prof. Emeritus Hans-Erik Pettersson, Prof. Hans Norin and Prof. Lotta Jakobsson for valuable comments on the earlier versions of the manuscript. Finally, we thank Kristina Mayberry (Mayberry Academic Services) for enhancing the language of the manuscript. We would also like to thank the anonymous reviewers for their dedicated review and valuable comments.

References

Ahlström, C., Victor, T., Wege, C., & Steinmetz, E. (2012). Processing of eye/head-tracking data in large-scale naturalistic driving data sets. *IEEE Transactions on Intelligent Transportation Systems, 13*, 553–564.

Baker, C. L., Jr, & Braddick, O. J. (1985). Eccentricity-dependent scaling of the limits for short-range apparent motion perception. *Vision Research, 25*, 803–812.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological, 57*, 289–300.

Björnstig, U., Björnstig, J., & Eriksson, A. (2008). Passenger car collision fatalities – with special emphasis on collisions with heavy vehicles. *Accident Analysis and Prevention, 40*, 158–166.

Broström, R., Ljung Aust, M., Wahlberg, L., & Källgren, L. (2013). What drives off-road glance durations during multitasking – capacity, practice or strategy? In: *3rd International conference on driver distraction and inattention*. Gothenburg, Sweden.

Craft, R. H., & Preslopsky, B. (2009). Driver distraction and inattention in the USA large truck and national motor vehicle crash causation studies. In: *First international conference on driver distraction and inattention*, September 28–29, 2009. Gothenburg, Sweden.

Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., et al. (2006). The 100-car naturalistic driving study, phase ii – results of the 100-car field experiment. In: *Report no: HS 810 593: NHTSA DOT*.

Dingus, T. A., Hulse, M. C., Antin, J. F., & Wierwille, W. W. (1989). Attentional demand requirements of an automobile moving-map navigation system. *Transportation Research Part A: General, 23*, 301–315.

Dozza, M., Moeschlin, F., & Léon-Cano, J. (2010). FOTware: A modular, customizable software for analysis of multiple-source field-operational-test data. In: *Second international symposium on naturalistic driving research*, Blacksburg.

Dukic, T., Hanson, L., & Falkmer, T. (2006). Effect of drivers' age and push button locations on visual time off road, steering wheel deviation and safety perception. *Ergonomics, 49*, 78–92.

Engström, J. A., Victor, T., & Markkula, G. (2013). Attention selection and multitasking in everyday driving: A conceptual model. In M. A. Regan, J. D. Lee, & T. Victor (Eds.). *Driver distraction and inattention, advances in research and countermeasures* (Vol. 1, pp. 27–54). Surrey, England: Ashgate.

Fitch, G. M., Soccolich, S. A., Guo, F., McClafferty, J., Fang, Y., Olson, R. L., et al. (2013). The impact of hand-held and hands-free cell phone use on driving performance and safety-critical event risk. DOT HS 811 757.

Fitch, G. M., Grove, K., Hanowski, R., & Perez, M. (2014a). Compensatory behavior of drivers when conversing on a cell phone: Investigation with naturalistic driving data. *Transportation Research Record: Journal of the Transportation Research Board*.

Fitch, G. M., Hanowski, R. J., & Guo, F. (2014b). The risk of a safety-critical event associated with mobile device use in specific driving contexts. *Traffic Injury Prevention*, 1–34. http://dx.doi.org/10.1080/15389588.2014.923566.

Funkhouser, D., & Sayer, J. (2012). Naturalistic census of cell phone use. In: *Proceedings of the 91st annual meeting of the transportation research board* (Vol. 2321, pp. 1–6), Washington, D.C.

Green, P. (2008). Driver interface safety and usability standards: An overview. In: *Driver distraction: Theory, effects and mitigation* (pp. 445–461), CRC Press.

Hickman, J. S., & Hanowski, R. J. (2012). An assessment of commercial motor vehicle driver distraction using naturalistic driving data. *Traffic Injury Prevention, 13*, 566–574.

Horrey, W. J., & Wickens, C. D. (2007). In-vehicle glance duration: Distributions, tails and a model of crash risk. *Transportation Research Record: Journal of the Transportation Research Board*, 22–28.

ISO. (2002). Road vehicles – Measurement of driver visual behaviour with respect to transport information and control systems. Part 1: Definitions and parameters. ISO 15007–1. In.

Kircher, K., Ahlström, C., & Kircher, A. (2009). Comparison of two eye-gaze based real-time driver distraction detection algorithms in a small-scale field operational test. In: *Driving assessment 2009: 5th international driving symposium on human factors in driving assessment, training and vehicle design*. Big Sky, Montana.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. In: *DOT HS 810 594*.

Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. *New England Journal of Medicine, 370*, 54–59.

Konstantopoulos, P., Chapman, P., & Crundall, D. (2010). Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accident Analysis and Prevention, 42*, 827–834.

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research, 25*, 296–324.

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

Land, M. F., & Horwood, J. (1995). Which parts of the road guide steering? *Nature, 377*, 339–340.

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

Lansdown, T. C. (2001). Causes, measures, and effects of driver visual workload. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, Workload, and Fatigue* (pp. 351–369). Mahwah, N.J.: Lawrence Erlbaum Associates.

Lappi, O., Lehtonen, E., Pekkanen, J., & Itkonen, T. (2013). Beyond the tangent point: Gaze targets in naturalistic driving. *Journal of Vision, 13*.

Lee, J. D., Moeckli, J., Brown, T. L., Roberts, S. C., Schwarz, C., Yekhshatyan, L., et al. (2013). Distraction detection and mitigation through driver feedback. Report No. DOT HS 811 547A, Washington, DC: National Highway Traffic Safety Administration.

Lehtonen, E., Lappi, O., & Summala, H. (2012). Anticipatory eye movements when approaching a curve on a rural road depend on working memory load. *Transportation Research Part F: Traffic Psychology and Behaviour, 15*, 369–377.

Liang, Y., & Lee, J. D. (2010). Combining cognitive and visual distraction: Less than the sum of its parts. *Accident Analysis and Prevention, 42*, 881–890.

Liang, Y., Lee, J., & Yekhshatyan, L. (2012). How dangerous is looking away from the road? Algorithms predict crash risk from glance patterns in naturalistic driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 54*, 1104–1116.

McCartt, A. T., Hellinga, L. A., Strouse, L. M., & Farmer, C. M. (2010). Long-term effects of handheld cell phone laws on driver handheld cell phone use. *Traffic Injury Prevention, 11*, 133–141.

Mourant, R. R., & Rockwell, T. H. (1970). Mapping eye-movement patterns to the visual scene in driving: An exploratory study. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 12*, 81–87.

NHTSA. (2013). Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices. In: *Docken No. NHTSA-2010-0053*. http://www.nhtsa.gov/staticfiles/nti/distracted_driving/pdf/distracted_guidelines-FR_04232013.pdf.

Olson, P. L., Battle, D. S., & Aoki, T. (1989). Driver eye fixations under different operating conditions. In: Ann Arbor, USA: The University of Michigan Transport Research Institute (UMTRI).

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009). Driver distraction in commercial vehicle operations, final report. In: Washington DC: Federal Motor Carrier Safety Administration.

Räsänen, M., & Summala, H. (2000). Car drivers' adjustments to cyclists at roundabouts. *Transportation Human Factors, 2*, 1–17.

Rockwell, T. H. (1988). Spare visual capacity in driving-revisited: new empirical results for an old idea. In: A. G. Gale, M. H. Freeman, C. M. Haslegrave, & P. Smith (Eds.), *Vision in vehicles II. Proceedings of the second international conference on vision in vehicles* (pp. 317–324). Nottingham, UK, 14–17 September 1987: Elsevier.

Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors, 48*, 362–380.

Sanchez, D., Garcia, E., Saez, M., Benmimoun, M., Pütz, A., Ljung Aust, M., et al. (2012). euroFOT D6.3 – Final results: User acceptance and user-related aspects. In (Vol. http://www.eurofot-ip.eu/en/library/deliverables/).

Serafin, C. (1994). Driver eye fixation on rural roads: insight into safe driving behavior. In: Ann Arbor, Michigan: The University of Michigan, Transport research Institute (UMTRI).

Sivak, M. (1998). The information that drivers use: Is it indeed 90 percent visual? *The UMTRI Research Review, 29*, 1.

Smith, D., Chang, J., Glassco, R., Foley, J., & Cohen, D. (2005). Methodology for capturing driver eye glance behavior during in-vehicle secondary tasks. *Transportation Research Record: Journal of the Transportation Research Board, 1937*, 61–65.

Stutts, J. C., Reinfurt, D. W., Staplin, L., & Rodgman, E. A. (2001). The role of driver distraction in traffic crashes. In: *AAA Foundation for Traffic Safety*. Retrieved Feb 2, 2012, from http://www.aaafoundation.org/pdf/distraction.pdf.

Summala, H., Nieminen, T., & Punto, M. (1996). Maintaining lane position with peripheral vision during in-vehicle tasks. *Human Factors, 38*, 442–451.

Summala, H., & Räsänen, M. (2000). Top-down and bottom-up processes in driver behavior at roundabouts and crossroads. *Transportation Human Factors, 2*, 29–37.

Swedish Official Statistics. (2013). Vägtrafikskador 2012, road traffic injuries 2012, Statistik 2013:9.

Tijerina, L., Barickman, F. S., & Mazzae, E. N. (2004). Driver eye glance behavior during car following. In: *DOT HS 809 723*.

Tsimhoni, O., & Green, P. (1999). Visual demand of driving curves as determined by visual occlusion. In: *Vision in vehicles 8*. Boston, MA.

Tsimhoni, O., & Green, P. (2001). Visual demand of driving and the execution of display-intensive in-vehicle tasks. In: *Proceedings of the human factors and ergonomics society*. *Annual Meeting*, 2, 1586.

Victor, T. (2010). The Victor and Larsson (2010) distraction detection algorithm and warning strategy, Volvo Technology.

Victor, T. W., Harbluk, J. L., & Engström, J. A. (2005). Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*, 167–190.

Wang, J. S., Knipling, R. R., & Goodman, M. J. (1996). The role of driver inattention in crashes; New statistics from the 1995 Crashworthiness Data System. In: *Association for the advancement of automotive medicine – 40th annual proceedings* (pp. 377–392).

Wierwille, W. W. (1993a). An initial model of visual sampling of in-car displays and controls. In A. G. Gale, I. D. Brown, C. M. Haslegrave, H. W. Kruysse, & S. P. Taylor (Eds.), *Vision in vehicles 4* (pp. 271–282). Amsterdam: North-Holland.

Wierwille, W. W. (1993b). Visual and manual demands of in-car controls and displays. In B. Peacock & W. Karwowsk (Eds.), *Automotive ergonomics* (pp. 299–320). London: Taylor and Francis.

Wierwille, W. W., & Tijerina, L. (1998). Modelling the relationship between driver in-vehicle visual demands and accident occurrence. *Vision in Vehicles, 6*, 233–243.

Wikman, A.-S., Nieminen, T., & Summala, H. (1998). Driving experience and time-sharing during in-car tasks on roads of different width. *Ergonomics, 41*, 358–372.

Zhang, H., Smith, M. R. H., & Witt, G. J. (2006). Identification of real-time diagnostic measures of visual distraction with an automatic eye-tracking system. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 48*, 805–821.

Zwahlen, H. T., Adams, C. C., & De Bald, D. P. (1988). Safety aspects of CRT touch panel controls in automobiles. In: A. G. Gale, M. H. Freeman, C. M. Haslegrave, P. Smith, & S. P. Taylor (Eds.), *Vision in vehicles II. Second international conference on vision in vehicles*, Nothingham, England (pp. 335–344). Amsterdam, North-Holland: Elsevier Science.