Holistic assessment of driver assistance systems: how can systems be assessed with respect to how they impact glance behaviour and collision

(article starts on next page)

Holistic assessment of driver assistance systems: how can systems be assessed with respect to how they impact glance behaviour and collision avoidance?

Jonas Bärgman[1] ✉, Trent Victor[1,2]

[1]*Division of Vehicle Safety, Chalmers University of Technology, Lindholmspiren 3, Gothenburg, Sweden*
[2]*Volvo Cars Safety Centre, Volvo Cars, Gothenburg, Sweden*
✉ *E-mail: jonas.bargman@chalmers.se*

Abstract: This study demonstrates the need for a holistic safety-impact assessment of an advanced driver assistance system (ADAS) and its effect on eye-glance behaviour. It implements a substantial incremental development of the what-if (counterfactual) simulation methodology, applied to rear-end crashes from the SHRP2 naturalistic driving data. This assessment combines (i) the impact of the change in drivers' off-road glance behaviour due to the presence of the ADAS, and (ii) the safety impact of the ADAS alone. The results illustrate how the safety benefit of forward collision warning and autonomous emergency braking, in combination with adaptive cruise control (ACC) and driver assist (DA) systems, may almost completely dominate the safety impact of the longer off-road glances that activated ACC and DA systems may induce. Further, this effect is shown to be robust to induced system failures. The accuracy of these results is tempered by outlined limitations, which future estimations will benefit from addressing. On the whole, this study is a further step towards a successively more accurate holistic risk assessment which includes driver behavioural responses such as off-road glances together with the safety effects provided by the ADAS.

1 Introduction

1.1 Background

There are several ways to avoid or mitigate injuries, and even reduce the number of crashes on our roads. Advanced driver assistance systems (ADASs; [1]) in vehicles typically aim to increase safety by avoiding crashes or mitigating their consequences by, for example, lowering the impact speeds. Actually, increased safety can be achieved by risk-, threat- and injury management. In risk management, ADASs support the driver to avoid threats, while in threat management, the critical situation has arisen and there is an ongoing threat that the ADAS supports the driver to avoid crashing or mitigate its consequences (e.g. lower the impact speed). Some ADASs are primarily for comfort or convenience, to reduce strain from driving. Such systems include adaptive cruise control (ACC, [2]), which keeps the headway to a lead-vehicle (LV) to a set value by controlling the vehicle throttle and brake; and lane keep assist (LKA; [3]), which supports lane keeping through steering control. Today there are also a variety of ADASs that combine ACC and lane-centring LKA [3], typically supporting the driver during highway driving. Hereafter this combination will be called driver assist (DA). Note that DAs and other convenience systems are also, implicitly or explicitly, risk management systems— as they are expected to increase safety—by maintaining better safety margins [2, 4]. A risk management system (e.g. DA) provides potential safety benefits from headway, speed and lane maintenance. When a traffic situation evolves into a critical situation despite this support, the threat management intervenes. Examples of threat management systems are forward collision warning (FCW; [5, 6]) and automatic emergency braking (AEB; [5–7]) systems. FCW warns the driver of an impending threat ahead in order to redirect the driver's gaze to the forward scene (the threat). FCW systems only warn (inform) the driver, and do not intervene by physically controlling the vehicle. AEB, on the other hand, intervenes automatically, braking the vehicle hard to avoid crashing with an object in front, or reducing the impact speed. The benefit of FCW and AEB has been shown in several studies. Cicchino [6] estimated that police

reported crashes are reduced by 27% with FCW and an additional 23% with AEB. Using a different study approach, Fildes *et al.* [8] showed a benefit of AEB of 38% crash reduction. Note, however, that these are estimates of benefit with systems that are early-on-market. More modern systems should have much higher effectiveness as technology is developing [9].

In manual-driving crashes (driving without any risk management systems that control the vehicle during normal driving), drivers looking (glancing) away from the forward roadway have, in a large number of studies, been shown to be a key crash causation mechanism – especially for rear-end crashes [10–12]. In these manual driving scenarios, the goal of FCW systems is to redirect the driver's gaze to the forward roadway [13, 14]. However, ACC (and the ACC component of DA) can also influence driver glance behaviour. Morando *et al.* [13] demonstrate that drivers typically shift their eyes back onto the forward roadway even before the FCW warning. This is an indication that the kinematic cue (deceleration) provided by the ACC prompts the driver to look forward—probably to check what is causing the unusual braking. In the present study, both the FCW warning and the kinematic cue from ACC/DA are modelled to redirect gaze towards the forward roadway [13].

Drivers glancing off-road are a key crash causation mechanism. As a result, there are government guidelines [15, 16] and other policy documents [17] that provide metrics for what is considered safe. For example, the national highway traffic safety administrations (NHTSA) distraction guidelines proscribe that glance durations while performing secondary activities are kept within 12 s total off-road glance time with glances primarily below 2 s individual glance duration. These guidelines can be traced back to early work in the glance-behaviour research domain by Rockwell [18], who presented a distribution of off-road glance durations associated with manual FM-radio tuning. The Rockwell distribution has been used extensively as a reference for glance behaviour that is considered safe enough (see review in [16, 19]). Therefore it is also included as a reference in the current study. In summary, all the available guidelines for 'safe task engagement' in vehicles are based on glance behaviours under manual driving

conditions (e.g. there were no ADAS systems in 1988). Consequently, when a specific driver-vehicle-interface (DVI) has failed to meet the guidelines, that statement is only related to what is considered safe in manual driving, without any ADAS systems.

Now consider risk management systems such as ACC and DA, where the vehicle partially controls the throttle and brake (and, for DA, steering as well). Note that the threat management systems FCW and AEB are always included in vehicles (packaged together) with ACC and DA. While using ACC and DA, some studies have shown glance behaviour changes, summarised here. Analysis of the large Field Operational Test EuroFOT [13, 20] showed that a percent road centre (PRC) reduction effect (few eyes on-path) occurs in steady-state driving with ACC, but that glances do return on-path in response to vehicle deceleration regulation if the lead vehicle brakes. In two studies of a large-scale naturalistic driving eye-glance FOT dataset, no striking differences in aggregate off-path glance duration distributions were found while using ACC compared to manual driving, rather the effect of ACC was characterised by longer glances on the path, yet there was a decrease in PRC (eyes on the path) [14]. Regarding visual time sharing, there was a slight tendency towards a higher median total task time with ACC active compared to manual driving [21]. Other studies have shown that the activation of risk management systems such as ACC and DA affects driver glance behaviour by producing glances off-road tending to be longer and the percent of time spent looking at the road tends to be less [22, 23].

For example, drivers performing a secondary task such as FM radio tuning may look at the radio longer when DA is active than during manual driving. Such a finding would suggest that glance behaviours from drivers performing secondary tasks like the standard FM radio tuning task with the DA active may risk failing to comply with the guidelines. The problem with such a scenario is that the DA system, together with the always-present FCW and AEB systems, may actually make the system significantly safer than the manual driving it is compared to and a holistic approach is needed to balance potential benefits and disbenefits. The ACC and DA's tacit influence on the driver's glance behaviour improve safety by reducing speeds, improving lane-keeping, and redirecting driver glances to threats [13].

In this study we demonstrate that assessing risk requires considering the combined impact on the safety of (potential) risks associated with glance behaviours induced by the presence of a risk management system and the (potential) benefits of such a system on safety associated with improved safety margin maintenance. There are, however, no studies that consider this combined effect; moreover, the current approaches (e.g. current NHTSA test procedure) for assessing glance-related risks cannot capture/assess this combined effect.

How can this combined impact on safety be evaluated? Actually, there have been several papers in the last few years that have used computer simulations to assess the impact of the driver glance behaviour [24–26], or the presence of ADAS [5, 27], separately, on safety. Thus, it is a natural step to assess the driver glance behaviour and the presence of ADAS jointly, using what-if, or counterfactual, computer simulations. This 'what if' naming reflects questions such as 'What if the car had an FCW and AEB instead?' and 'What if the driver would have glanced off-road in this way instead?'. The what-if simulation is assessed using existing crashes or near-crashes, such as those collected from naturalistic driving data [24, 25], event data recorders [28], crash reconstruction [29, 30], and even, time-series data from everyday driving [25, 31].

Although some may argue that using computer simulations for safety assessment is less valid, than, for example, using controlled experiments but limitations exist with both methods. To address the validity of using simulations, it is useful to consider the state-of-the-art in the domain of passive safety (injury management, protecting the occupants of the car from injury and death after a crash; [32, 33]). Computer simulations are currently used to evaluate a wide range of functionality from individual vehicle components to complete vehicle structures [34, 35]. Although the use of computer simulations to aid the product design process in the passive-safety domain is more mature than their use in ADAS

design and evaluation, evaluations of the impact of ADAS on driver pre-crash behaviour is quickly developing. Analogous to development of human body models (HBM; [36]) – computer modelling of the (physical) human for use in crash simulations – so too can we expect the development of driver models in what-if crash-avoidance simulations.

A recent position paper [37] on human factors for automated vehicles from the EU project Coordination of Automated Road Transport Deployment for Europe (CARTRE) identifies simulations as an ideal for the assessment of vehicle automation. Also, Bellet *et al.* [38] contributes to the virtual assessment of vehicle automation by the outlining a high-level framework that combine models of driver's cognition with ADAS and automated vehicles for holistic assessment.

In sum, as computer simulations incrementally improve, their validity improves through calibration with results from experiments, naturalistic data, and crash databases. The use of simulation is particularly essential when there are few, or even no, other available methods to assess an effect; this, we argue, is the case for the combined effect of driver glance behaviour and the presence of an ADAS.

1.2 Aim

This study has two aims: (i) to demonstrate the need for a holistic safety-impact assessment of ADAS and (ii) to propose a method for carrying out this assessment. 'Holistic safety impact' here means the combination of (i) the impact of the change in drivers' off-road glance behaviour due to the presence of the ADAS on, with (ii) the safety impact of the ADAS alone (e.g. by autobraking).

To accomplish these aims, this study demonstrates the estimated combined impact on crash risk of a set of ADAS (DA, ACC, FCW, and AEB), with the corresponding off-road glance-distribution changes due to the presence of those systems.

2 Method and materials

This study used the same dataset and fundamental simulation method as in [26]. The following sections summarise the approach, providing the most important details; for further information, such as signal processing, crash selection, data reduction, and detailed FCW and AEB algorithm implementations, see [26]. After a description of the data used and a brief description of the ADAS algorithms used are provided, the what-if simulation approach is explained and the performed simulations and subsequent analysis are presented.

2.1 Data

The data used come from the second strategic highway research program (SHRP2), the world's largest in-vehicle naturalistic driving study to date [39]. In SHRP2 3247 primary participants across six sites in the US drove instrumented vehicles a total of almost 80 million kilometres over two years, as part of their everyday lives. In addition to everyday uneventful driving, almost 7000 near-crashes and more than 1000 crashes (severity level 1–3, see [39]) were captured in the recordings. In this study, a subset of the rear-end crashes from SHRP2 was used, consisting of 34 crashes in which the SHRP2-instrumented car struck a lead vehicle (see [26] for details). These crashes were a subset of the 46 crashes used in [11]. The twelve excluded crashes had either missing or corrupt speed or acceleration data, or the relative acceleration between the two vehicles was unrealistic (see [24]). Note that due to missing or poor-quality radar data for several events in the original SHRP2 dataset [11], manual annotations (performed as part of [11]) was used to provide data about the relative kinematics between the two crashing vehicles in the seconds leading up to the crash. Approximately 12 seconds of time-series data was available prior to the crash point for this study. Ten Hz time-series data included the following vehicle (FV) speed and acceleration, the LV speed and acceleration (relative to the FV), and the distance between the two vehicles. Based on these data, theta (the vertical angle formed by the left and right edges of the LV from the
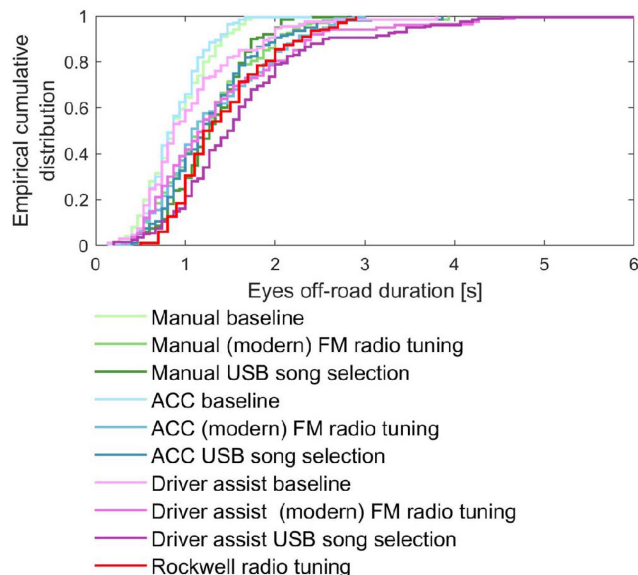
Fig. 1 *Cumulative glance off-road distributions from the aggregation of 20 drivers' behaviours under nine experimental conditions, with the Rockwell [18] radio tuning task as a reference*
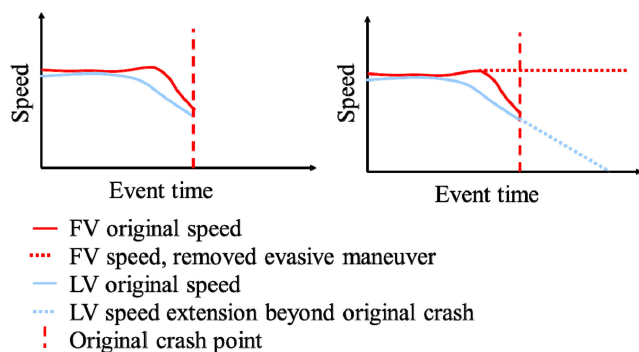

Fig. 2 *Illustration of the original kinematics (left), and seed crash (right) with the FV's evasive manoeuvre removed, and the LV speed extrapolated beyond the crash*

perspective of the FV driver; [40]), and theta-dot (looming: the derivative of theta; [40]), were derived. These data for each of the 34 SHRP2 crashes were used to represent original crash kinematics in the what-if simulations performed in the study.

In the what-if simulations, two additional variables were used: distribution of the maximum driver deceleration and eyes-off-road glance distributions (EOFF). The distribution of the maximum driver decelerations across crashes was the same in this study as in [26] (Appendix A, Figure A3). The EOFF distributions were obtained from an on-road experiment (20 drivers drove under nine conditions; further details to follow) and the aforementioned Rockwell radio-tuning task (as a task used as a reference for 'acceptable risk'; [19]). Also, for comparison, Fig. 1 shows the cumulative off-road glance distributions of these ten off-road glance distributions (all drivers in the study pooled). The off-road glances of manual driving are below the Rockwell (cumulative) reference distribution (across tasks), while the off-road glances for all tasks in the ACC and DA driving condition in part have longer off-road glances than the reference distribution. Within each driving condition, a baseline has shorter off-road glances than FM radio tuning, which in turn has shorter off-road glances than USB song selection. These cumulative distributions are in line with previous findings (see [13, 14, 16, 21]).

2.2 On-road experiment

Twenty participants (nine women and 11 men) aged between 27 and 62 drove a Volvo MY2016 XC90 on the main highway in the Göteborg area. Drivers were recruited through email to Volvo employees. The selected participants could not be involved in

vehicle development, could not work as test drivers, had not participated in similar studies before, and had a minimum driving experience of at least 5000 km during the previous year. The data collection was performed independent of the current study (as part of another study). The current study was designed and executed after the completion of the data collection, reusing existing data.

All participants were first given information about the study, and after agreeing to participate, signed informed consent and filled in a demographics questionnaire. Once in the car, they underwent a learning session, as described in 'Test participant training recommendations' in the NHTSA-guidelines [16], in order to familiarise them with the secondary tasks and the specific way those tasks should be performed. Participants were also introduced to the ADAS-systems, if needed. They then drove the car with the test leader, to get a first-hand feel of the car and systems in traffic. A test-leader manual was used to ensure that all participants experienced the same procedure.

When the participants were comfortable with the setup, the eye tracker (Smart Eye Pro 6.0) calibration was performed and the participants had the opportunity to repeat the tasks a final time with the test leader present. Then the participant drove alone from the preparation location towards the starting point on the E6 highway from Gothenburg to the city of Kungälv. The drivers were instructed to drive to Kungälv and back on the same road.

While driving, an indication of which task was next to be performed was read by the participant on an iPad (e.g. 'USB task now'). They informed the test participants which route to drive, which ADAS modality to drive in, and which secondary tasks to perform. The three ADAS modalities were: manual driving, driving with ACC, and driving with a DA (in this case the Volvo Pilot Assist system). After the iPad instructed participants which ADAS modality to use, the participants themselves chose when to step to the next instructions on the iPad: that is, initiation and performance of the secondary task were self-paced. After all tasks had been completed for an ADAS modality, the participant self-rated driving performance on a 1–7 scale on the iPad, and answered the three SAM-scale ratings [41] (part of the original study, for which the data collection was designed – these were not used in the current study).

The radio task involved changing from one radio station to another with a manual tuning slide bar. USB song selection involved finding and selecting a song from a stored list.

A forty-second long data segment, starting 30 seconds after the completion of the last task in each ADAS modality, was used as a baseline for that modality.

2.3 What-if simulation design

2.3.1 What-if simulation overview:
What-if simulations can be performed in a variety of ways. This study used the same approach as in [26], consisting of six steps (the steps are described in more detail later in this section). First, the crashes to be used as original crashes are identified. These should have kinematic time-series data (speed and acceleration) of both the FV and the LV, as well as the range between the two, up until the crash point. Second, for each original crash, any evasive manoeuvre made by the FV is removed, and the LV kinematics are extrapolated beyond the crash point (see Fig. 2). These crashes, now called seed crashes, correspond to what the original crash would have been if the FV driver had not performed any avoidance manoeuvre. Third, (in this study) 200 samples are drawn from the EOFF glance-duration distribution for the glance behaviour being studied (e.g. the behaviour associated with DA FM radio tuning; Fig. 1). Fourth, a model is used to simulate counterfactual ('what if the driver glanced away here') off-road glances; it positions the sampled (what-if) off-road glance in the time-series according to a glance placement model. The output from this step is a time-series of an event as if it was collected in an experiment, with the kinematics of the FV and the LV, and the (last) glance behaviour of the driver of the FV. In an experiment, the drivers' off-road glance timing and duration can be captured with an eye-tracker system, while here the combination of the kinematics and the off-road glances is based on the stochastic selection of off-road glance durations, and an off-
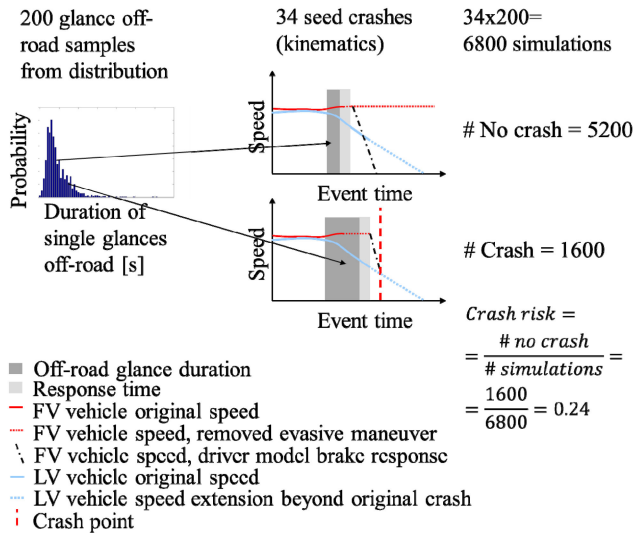
Fig. 3 *Representation of the what-if simulation process for assessing the impact of drivers' off-road glance distributions on crash risk*

road glance placement model (wherein the time-series the off-road glance should be placed, see 2.3.3 and [24, 26] for further description). Fifth, what-if simulations are performed for each combination of seed crash and off-road glance sample, using a model of the driver's response process to a critical event. The response-process model for the driver includes the timing of braking initiation, the jerk, and the maximum deceleration. Each what-if simulation will either be a crash or a near-crash (Fig. 3). Sixth, and finally, the proportion of crashes out of the total number of simulations is calculated: this is the crash risk metric specific to the off-road glance distribution being studied. Using this approach, different off-road glance distributions can be compared by comparing crash risk estimates (proportions of crashes in the simulations) between the distributions.

The addition of ADAS systems in the simulation process model means that the combined effect of the glance and the ADAS can be evaluated. In this combined model, the FCW, ACC, and DA systems redirect the driver's off-road glance (the glance sample placed in the time-series) in different ways, according to their respective glance-redirect model. Then the driver-response process model is activated, so braking occurs when the model indicates the driver would have braked. For AEB, when the AEB-algorithm trigger threshold has been reached, the deceleration associated with the AEB is applied to the FV. It is thus possible to compare the various combinations (of off-road glance behaviours and ADAS) with respect to estimated crash risk. The following sections describe this study's what-if simulation components in more detail.

2.3.2 Creating seed crashes: The 34 original crashes were identified and converted into seed crashes by removing the evasive manoeuvre for the FV. The speed of the LV was extrapolated beyond the crash point (Fig. 2), so that simulated driver and ADAS brake responses would not be influenced by the driver's response in the original crash. Seed crashes have the same initial kinematics as the original crashes without the original FV responses to the critical event (see [26], for details).

2.3.3 Applying off-road glances to the seed crashes: For quality assurance, all glance time-series produced by the eye-tracker were manually inspected for quality and when the eye-tracking quality was not sufficient, the gaps were manually annotated using the following general rules: (i) driver gaze was annotated from the start of the task to the end of the task, (ii) transitions were coded as part of the previous glance, (iii) blinks were not taken into account in the annotations (if a test participant moved their gaze off path during a blink, the glance was annotated as off-road from the point in time when the test participant's eyes were closed), and (iv) whether a test participant was looking on or

off-path was subjectively determined through video data of the drivers' eyes.

After quality assurance, samples from the off-road glance distributions were applied to the seed crashes (Fig. 3). Two hundred EOFF glance durations were sampled, using a Monte Carlo sampling approach, from the distribution being evaluated (each of ten different glance behaviours used in this study, respectively; see Fig. 1). Each sampled off-road glance was applied to every one of the 34 seed crashes, creating a total of 6800 simulations. 'Applying' a glance means placing it in the time-series, simulating an off-road glance by the driver during that time —instead of replicating what the driver actually did in the original crash event.

For the simulation to be valid, it is necessary to place the off-road glance at a realistic time point in the time-series data. From the few models which do this (e.g. [24, 25]), we chose the same glance placement model as in [26], model F. This glance placement model is based on studies addressing the question: 'After how much looming (what value for thetadot) do drivers start reacting to critical situations?' The model assumes that there is an equal probability of initiating/not initiating a task when thetadot is lower than 0.01 rad/s; however, when thetadot is greater than 0.01 rad/s, drivers would not initiate any task (or perform any off-road glance). Thus, according to this model, only glances that overlap thetadot equal to 0.01 rad/s matter, and placement of the glance in the time series is then conditioned by this overlap.

Glances overlapping thetadot = 0.01 rad/s were simulated with the process described here. First, a sample (T_{EOFF}; one of the 200) was drawn from the eyes-off-road glance distribution being studied. Second, a random value from the range [0, T_{EOFF}] was drawn. This duration was then placed in the kinematic time-series, starting at the point in time where thetadot reached 0.01 rad/s. This simulates equal probability of a glance overlapping thetadot 0.01 rad/s during T_{EOFF} (note that there are statistical techniques that facilitate the overlapping approach analytically (see [24] for one such implementation), but here the pragmatic empirical sampling approach was used for clarity).

2.3.4 Driver response model: Like the glance-placement model, the driver-response model was the same one used in [26], model F. A brief overview of this response model follows.

As described previously, in this study we assume that there is an equal probability of glancing away before thetadot = 0.01 rad/s— the threshold based on [42], but when this threshold is reached the driver never looks away and reacts as fast as possible. The response timing is thus kinematically dependent. Further, the time to respond to the visual stimuli of the LV looming was set to 0.5 s, also based on [42]. The driver response model must, for completeness, consider two cases. In the first, the driver is looking at the road when thetadot reaches 0.01 rad/s, and the model assumes that the driver starts braking 0.5 s after the threshold is reached. However, since this implementation ignores on-road glances, this case is not relevant. In the second case, which is the one considered for this study, when there is an EOFF overlapping thetadot = 0.01 rad/s, it is assumed that the driver starts braking 0.5 s after having looked back on the road.

The driver response model handled FCW warnings in the following way: If drivers were gazing off-road at the time of the warning, they were assumed to look back on the road instantaneously—after which the response model described above took over ('if thetadot>0.01 rad/s, start braking 0.5 s after'). If the drivers' eyes were on the road at the time of the warning, the original response model was used (not relevant for the current implementation).

In this study it is assumed that the driver's only avoidance manoeuvre is braking (not steering). The deceleration profile is based on the kinematically dependent jerk (deceleration derivative) found in [42], and a sample of from the maximum deceleration distribution [26]. A sample each of the 200 simulations of eyes-off-road glances (see above).

			ADAS-modalities in the what-if simulations			
			Manual (no-system)	Manual (+FCW+AEB)	ACC (+FCW+AEB)	Driver Assist (+FCW+AEB)
Reference	Manual	Rockwell				
Off-road glance distributions from on-road experiment	Manual (+FCW+AEB)	Baseline				
		FM tuning		Factual		
		Song selection				
	ACC (+FCW+AEB)	Baseline	///			
		FM tuning	///		Factual	
		Song selection	///			
	Driver Assist (+FCW+AEB)	Baseline	///			
		FM tuning	///			Factual
		Song selection	///			

Factual - same configuration in on-road experiment and what-if simulation
Non-existent - glance behavior and ADAS configuration not in experiment
/// Traditional glance behaviors assessment - Non-existent and unrealistic

Fig. 4 *Conceptual table of 40 what-if simulation conditions, combining real off-road glance distributions (rows) and ADAS simulations that are either factual (an ADAS modality in the on-road experiment) or non-existent (a glance/ADAS-modality combination not a part of the on-road experiment)*

2.3.5 Automatic emergency braking: The AEB algorithm is the same one as in [26], which is based on the threat-assessment AEB algorithm described in [7]. The AEB algorithm used here takes the current (instantaneous) FV and LV speeds and accelerations, together with the distance between the two vehicles, to model the future states of both vehicles. The deceleration required to stop the FV before it crashes into the LV is calculated for each time point and compared to a threshold (AEB algorithm parameter; here 7 m/s^2). If the threshold is reached, the AEB is activated and the vehicle brakes.

2.3.6 Forward collision warning: The FCW algorithm, too, is the same as in [26], which in turn is based on the threat assessment AEB algorithm described in [7]. The only two differences between the AEB and FCW models are the threshold value (2.3 m/s^2 for FCW and 7 m/s^2 for AEB), and the fact that the AEB model simulated a braking intervention and the FCW model only simulated a warning to the driver.

2.3.7 Adaptive cruise control and DA: Neither ACC nor DA were modelled (implemented) to control the vehicle at any time (neither braking nor steering), but both were modelled to redirect glances to the forward roadway. That is, based on [13], when ACC or DA was active, it was assumed that the driver did not look off-road after TTC^{-1} 0.21 s^{-1} (i.e. if the driver had the eyes off the road when TTC^{-1} reached 0.21 s^{-1}, glances were assumed to be on-road after that). Morando *et al.* [13] demonstrate that the kinematic cues of ACC typically redirect driver glances back to the roadway no later than (approximately) $TTC^{-1} = 0.21$ s^{-1}.

2.4 Performed what-if simulations

In this study, two sets of what-if simulations were performed. The aim of the first set of simulations was to demonstrate the need to consider off-road glance behaviours and ADAS systems in combination when assessing a specific task or driver-vehicle interface, rather than the off-road glance behaviours alone. To create the simulations, combinations of the four different ADAS modalities and the ten different driver off-road glance distributions were what-if simulated (40 conditions; Fig. 4). The four modalities were manual driving (no system), FCW and AEB in manual driving, adaptive cruise control with FCW and AEB, and DA with

FCW and AEB, respectively (columns in Fig. 4). The ten driver glance behaviours comprised baseline and two secondary tasks (rows in Table 1) for each of the ADAS modalities (including the ADAS modality manual driving (no system) which was not included in the on-road experiment). As noted, the baseline for each modality was the 40 s of no-task driving after the end of the last task for that modality in the experiment. The two tasks studied here were (modern, touch-screen based) FM-tuning, and USB song selection (additional tasks were analysed in the study, but results from those are not reported here).

Out of these 40 what-if simulations, only three (marked in dark grey as Factual in Fig. 4) were estimating the actual risk associated with the specific driving conditions the drivers were exposed to. That is, these three simulations used the glance behaviours that were recorded in the on-road experiment in each of the ADAS modalities. In addition, 37 simulations were performed (Fig. 4) where there was no corresponding glance/ADAS-modality combination in the on-road experiment. Such simulations are hereafter called non-existent. For example, the top row in Fig. 4 describes the simulations combining glance behaviours from the Rockwell [18] task with the four ADAS modalities. These simulations evaluate the risks of driving in the four ADAS modalities, had the driver's off-road glances been exactly like those in the Rockwell [18] study. Similarly, column one in Fig. 4 describes simulations in which the driver performed off-road glances as in each of the ten tasks/baselines without an ADAS. In this study we argue that this column's (manual driving without any ADAS) simulations are particularly unrealistic, especially the estimated risks associated with the glance behaviours that are directly influenced by the presence of active support systems (ACC and DA, the bottom six rows in the no system column in Fig. 4). Those simulations ignore the fact that those glance behaviours are directly related in response to systems that also may provide a safety benefit and therefore should be considered in a holistic manner.

The results from the simulations in the twelve cells of FM tuning in combination with the four ADAS modalities (Fig. 4) are presented as box plots, together with their respective means. The estimation of risk for the Rockwell glance behaviour without any ADAS (leftmost top cell in Fig. 4) is also included in all boxplots, as a reference.

The results from the first set of simulations were analysed from several different perspectives—primarily based on the box-plots of the what-if estimated risks across drivers for different combinations of the 40 cells, but student *t*-tests were also used. A conservative alpha of 0.001 was used in the tests to avoid issues with multiple testing.

The aim of the second set of simulations was to demonstrate the impact of different proportions of complete ADAS system failures on the crash risks, addressing the question 'How risky is it for the system to fail in X% of the critical situations assessed through what-if simulations?' where X is 0 to 100% in steps of 10%. To accomplish this, the simulations in the rightmost bottom cell in Fig. 4 is used as a basis. Those simulations estimated the risk of the combination of the factual off-road glance distribution produced in the experiment when the DA system (with FCW and AEB present) was active. By randomly disabling the ADAS in a range of different proportions (0–100%) of simulations (of the rightmost bottom cell in Fig. 4), the impact of system failures on safety was assessed. Again, the primary assessment was performed by visualising, describing, and discussing the results represented by boxplots.

Note that this paper aims to present and discuss a method concept, rather than make absolute comparisons of risk.

3 Results

Fig. 5 shows the simulation of (modern) FM tuning in all ADAS modalities in Fig. 4. The manual (no system) modality (column) in Fig. 5 corresponds to the traditional assessment of off-road glance duration changes. It is of particular importance to compare the results of these non-existent combinations of manual (No system) estimates with the factual simulations of ACC(+ FCW + AEB) and
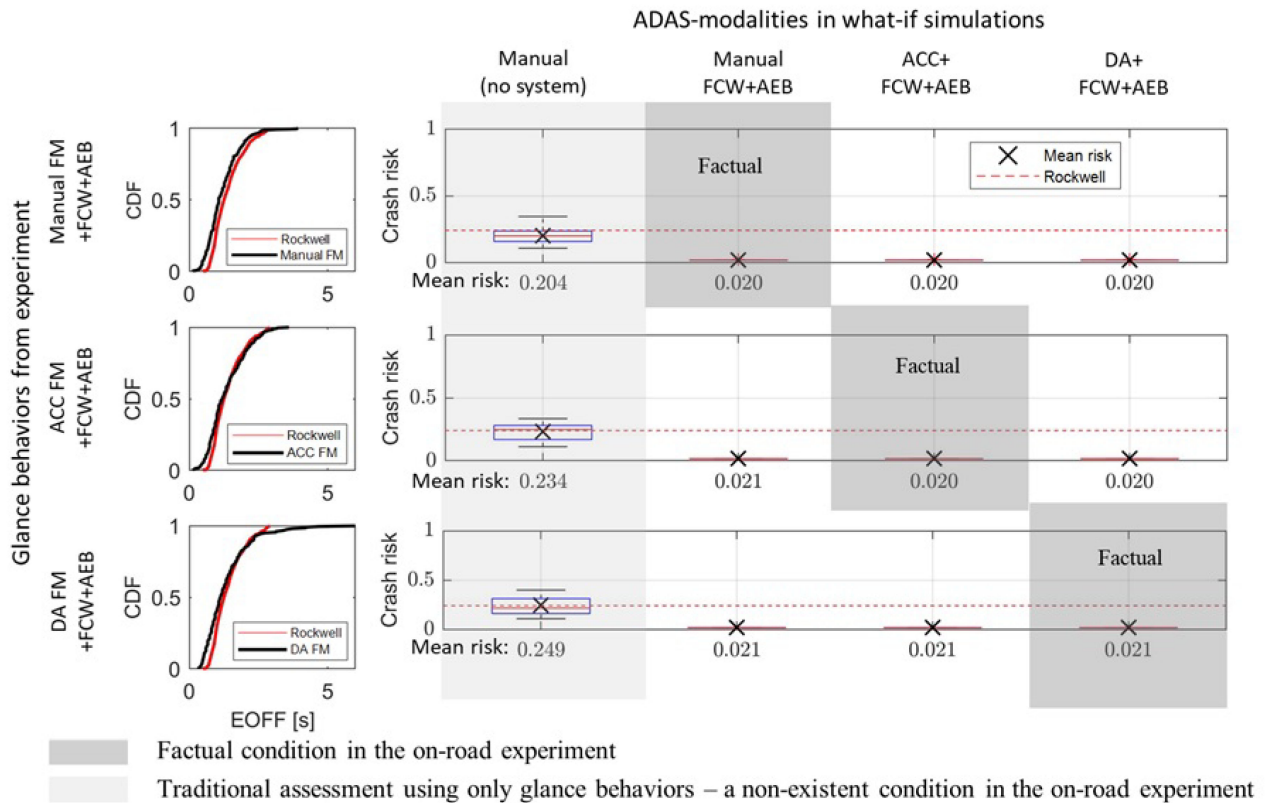
Fig. 5 *Left panels show the cumulative distributions, across all drivers in the experiment, of off-road glances for (modern) radio tuning, for each of the three automation modalities (top, manual; middle, ACC driving; bottom, driving with a DA system). The Rockwell radio tuning task off-road glance distribution is included as a reference in all three left panels. The right panels show the crash risk results of the factual (dark grey boxes) and non-existent what-if simulations for FM-tuning as boxplots, corresponding to all FM tuning task cells in Table 1. The boxplots represent distributions of crash risk across drivers. That is, the crash risks have been estimated (simulated) for each individual driver individually – drivers were then combined in the boxplot. The crash risk associated with all drivers performing the radio task, as documented in Rockwell [18] is shown as a reference*
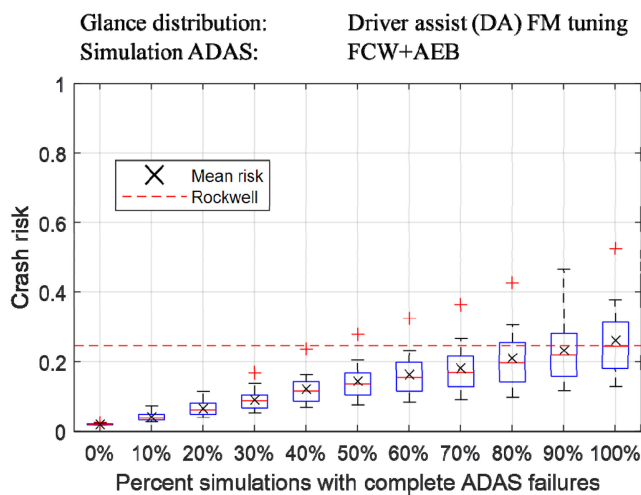

Fig. 6 *Results from simulations where FCW + AEB were completely disabled in zero to 100 percent (x axis) of the 6800 simulated events. The boxplot represents the distribution of crash risks across the 20 drivers. The risk estimates of manual (no system) driving with Rockwell is shown as a reference*

DA(+ FCW + AEB), respectively. The factual risk estimations show a large reduction compared to the risk estimation for the non-existent manual (no system) condition for each row – corresponding to the ADAS safety efficiency in the simulation. Actually, as soon as FCW + AEB are introduced, crash risks are drastically reduced (compare the leftmost column of the right panel with the three rightmost columns; $t(36)>11$, $p < 1 \cdot 10^{-9}$, when comparing simulations with manual (no system) and simulations that include FCW + AEB).

However, there is no significant crash risk reduction between FCW + AEB what-if simulations (second column with boxplots in Fig. 5), and the risks associated with ACC and DA with FCW + AEB (third and fourth column of box plots in Fig. 5), respectively. Similarly, there is no significant crash risk increase with an increase in the level of automation, and all simulations with at least FCW + AEB are far below the Rockwell crash risk.

Fig. 6 demonstrates how complete (and silent) failures of ADAS can affect the mean crash risk. These failures are on-top of the system limitations for the specific implementation. The figure shows the what-if simulated risks of FM radio tuning while the vehicle is operating in the DA automation mode, with different proportions of complete FCW + AEB failures. The specific implementation of the (ideal) FCW + AEB used in this study reduced the mean crash risk from 100% system failure (no system) to zero percent failures (working optimally as implemented) by more than a factor of 10 (>90% reduction). Fig. 6 shows an almost linear reduction in the effectiveness of FCW + AEB between zero and 100% failures. Similarly, there is an almost linear relationship between the quartiles (and even the outliers) from no failures to complete failure. Although with 100% failures there is a slightly higher risk than with Rockwell, when there is a 90% failure rate the crash risk is already lower than that of Rockwell. At 60% failures, all drivers except the outlier still have a crash risk below the average Rockwell risk.

Between-driver variability varies across tasks and ADAS modality. For simulations with no system (Fig. 7), the difference in the standard deviation of crash risk is similar within tasks between baseline and ACC, while the standard deviation of crash risks of DA is higher than for both baseline and ACC, for each task. For DA(FCW + AEB) simulations (Fig. 7), this effect of driver variability is basically eradicated.

Further, song selection in DA driving has a wider interquartile range and higher-risk outliers than manual song selection. The means of the tasks with larger variability are naturally higher, as
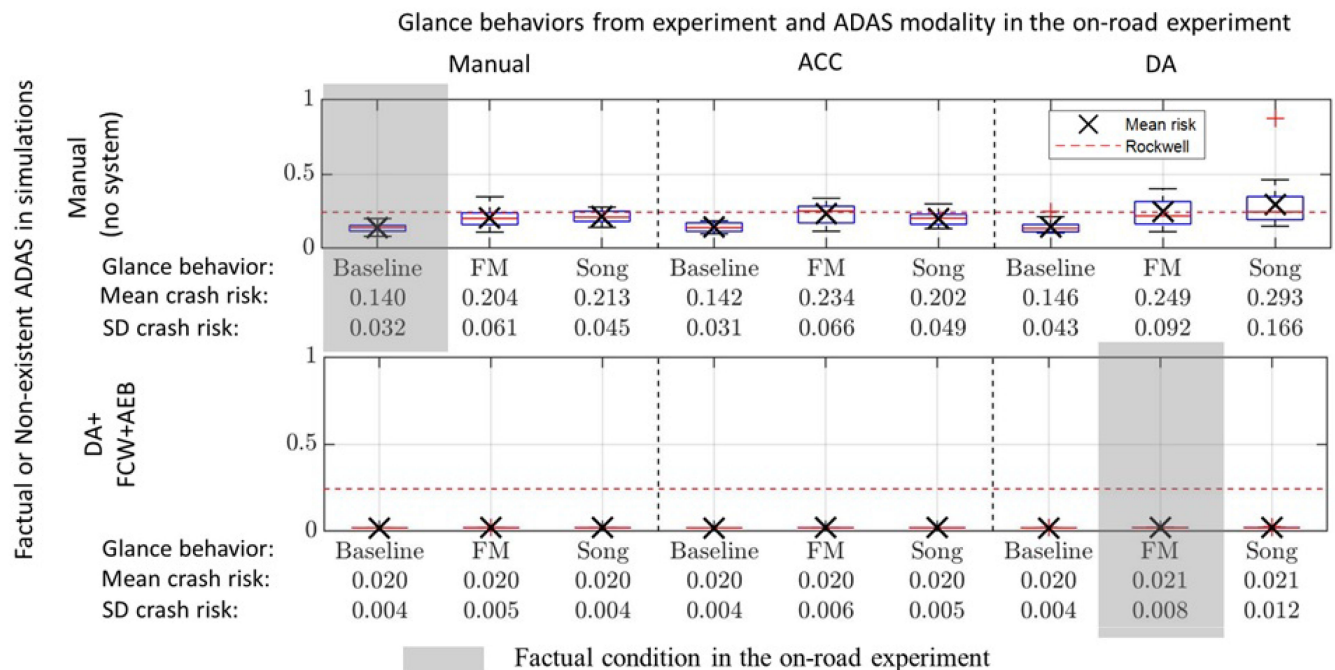
Glance behaviors from experiment and ADAS modality in the on-road experiment

Fig. 7 *These graphs are the results from simulations of the leftmost and the rightmost columns in Figure 4 (as if transposed). A comparison of crash risk estimation across driver glance-off road distributions of different ADAS modalities, but with the non-existent manual (no system) in the simulations. In addition, the same is performed with the DA with FCW and AEB systems in the simulations instead. Only the simulation in grey is factual*

risks are constrained in the downward direction (risk cannot be lower than zero) but not upwards (higher risk) in any practical sense. The mean crash risks for the modern FM tuning task and the Rockwell tuning task are similar, although the modern FM tuning task may be slightly less risky even with 100% failure rate (recall that this is a condition that could not reasonably even occur as glance changes associated with DA would not occur if the system is not active).

4 Discussion

This study develops a method for holistic risk assessment focusing on the combination of glance behaviour and ADAS safety effects. It uses what-if computer simulations to (first paper aim) demonstrate the need for the combined risk assessment of (i) the effect of off-road glances, and (ii) the effect of ADAS systems (FCW, AEB, ACC, DA) being present and active when the glance data was collected. That is, we argue the proposed method and its future developments are a more accurate way to balance the effects of glance behaviour and ADAS in a risk assessment (second paper aim). It follows that making conclusions about risks related to glance behaviour changes independently of consideration to larger potential benefits of ADAS would be much less accurate (as indicated by our results here e.g. Figs. 6–8). What-if simulations can assess such holistic safety effects and there are other ongoing efforts in to pursue assessing the combination of driver cognitive influences (e.g. precautionary behaviours, including glance behaviours and decisions like speed reduction decisions) and ADAS [38].

4.1 Holistic risk assessment

Longer off-path glances can be understood and modelled, within the cognitive neuroscientific predictive processing framework, as an understandable consequence of a perceived reduction in need to update that the predictive (generative) model is proceeding according to the plan (see [43, 44]).

The cumulative off-road glance distributions across tasks and ADAS modalities in Fig. 1 can be compared with the risk estimates from what-if simulations in the top panel of Fig. 7. The second boxplot from the left in the top panel (manual no system row) shows that the FM radio tuning results are comparable to the Rockwell reference task but includes a variance (a similar variance would have been present in the Rockwell task if they would have

published variance). In strong contrast to the crash risk for FM, song and baseline (bottom panel, Fig. 7) are significantly reduced to a level much below the manual baseline risk level (top panel, Fig. 7), across all ADAS modalities ($t(34)>15$, $p<1\cdot10^{-9}$, for all four tests).

The results of the simulations performed in this study strongly support the argument that potential risks related to glance behaviour changes need to be assessed together with the larger safety benefits from ADASs (FCW, AEB, ACC, and DA). If a holistic safety assessment is not considered, it may lead to a missed opportunity to realise the safety benefits of the ADAS. This may, in turn, potentially lead to the discouragement of the implementation of safety technology if glance behaviour changes are considered in isolation.

In contrast to the method we propose, current glance behaviour assessment methods related to in-vehicle interface interactions do not consider ADAS effects. For example, the NHTSA distraction guideline [16] and others [19, 45] are typically based on two main types of studies: (i) controlled experiments in driving simulators, on test-tracks, or on-road (see [16] for details), and (ii) studies of glance behaviour and their relation to the risk of being in a crash or near-crash [10, 46]. What is common among the studies on which current guidelines are based (as far as we have been able to establish) is that they have not addressed the benefits of either risk- or threat-management ADAS. For example, when the Rockwell [18] radio task is used as a basis for a distraction guideline, it ignores the presence of any ADAS in the vehicle—even if the ADAS being assessed both induces longer glances and improves safety (e.g. redirects glances).

As FCW and AEB are practically always present when ACC or a DA system is active, comparisons with and without ACC/DA should naturally be made with FCW and AEB present; the net risk change of adding ACC or DA to the FCW and AEB should be evaluated. In our results, the positive safety impact of FCW/AEB is much larger than the effect of ACC/DA. Further, effects of longer glances associated with ACC or DA are overwhelmed by the positive effects of the threat management systems FCW and AEB.

One argument for the current-practice 'no system' assessment, which is sometimes raised when the necessity of a holistic approach is presented, is that there may be situations where ADAS fail. However, in this study we demonstrate (Fig. 6) that what-if simulations can be used to investigate at what proportion of ADAS failures in critical events the risks start to approach the risk of some

1064

IET Intell. Transp. Syst., 2020, Vol. 14 Iss. 9, pp. 1058-1067

reference task/threshold. It is here important to note that it is highly unlikely that ADAS failures (for example, ACC or AD) would only occur in critical situations (if the systems fail in critical situations, it is likely that the system is unreliable also in everyday (normal) driving). If it is unreliable, it is unlikely that the driver would continue to feel safe enough to look away longer; he or she would likely turn the system(s) off and be considerably more attentive. This is predicted from and modelled in the predictive processing framework [43]. The likely result would in any case be that the drivers change back to manual-driving-like glance behaviours. Consequently, if a risk management driver support system (e.g. ACC or DA) has any risk-reduction properties, it is unlikely that ADAS failures would increase the risk associated with ADAS induced glances, since drivers would change back to manual-driving like glance behaviours.

4.2 Limitations

This study represents an incremental development of the methodology towards a successively more accurate holistic risk assessment which includes driver behavioural responses such as glances and the safety effects of the ADAS. Previous papers [24, 26] have taken the first steps. Important limitations likely affect the accuracy, validity, and generalisability of the risk estimations to absolute crash-risk. Successive improvements will be achieved by addressing limitations, as outlined below.

Firstly, the small sample (34 crashes) is not likely to be representative of (and thus not be generalisable to) the population level (here, rear-end crashes in the US). Interpretations of the actual absolute estimated risks through the what-if simulations should be treated with care, if made at all. Future incremental steps should include a more representative sample.

Secondly, the FCW and AEB used in the simulations are ideal systems, overperforming compared to the benefits of those systems in the field according to post-hoc studies. Reasons for the over performance include that no sensor degradations are included in the FCW and AEB models, and that the algorithms assume pure rear-end situations (with 100% overlap between the following and the lead vehicles). The latter reason is particularly important, as FCW and AEB algorithms typically generate warnings or interventions later when this overlap is smaller, as the driver may avoid the situation by small manual steering interventions. This delay avoids nuisance warnings and interventions. The combined model of FCW and AEB that was used likely overestimates risks by a factor of two, approximately (compare with [24], as well as [6, 8, 29]). However, even if the FCW and AEB benefits are only half (in line with [6, 8]) of the benefits estimated in this study's what-if simulations, the conclusions of this paper stand. Consider, e.g. Fig. 6, reducing the benefit to half allows ∼30% complete system failures (on top of the reductions due to sensors etc., present in real traffic today), for 75% of the drivers to be safer than they would be having performed Rockwell radio tuning glances during manual driving. Future studies should incorporate sensor limitations and scenario constraints into their ADAS models—preferably models from ADAS production.

Thirdly, the ACC and DA simulations are conservative; they are likely to produce more crashes than would occur in real life in two ways. First, the time headway (THW) was not modified (increased) from the original SHRP2 crash in any of the crashes with ACC or DA active. Having a very short THW with ACC or DA active is unrealistic, as an active ACC or DA would have increased the THW, reducing the risk of a crash. Future studies may choose to investigate the THW increases resulting from active ACC or DA systems. Including these increases in this study would likely have strengthened the main conclusions of the paper – the current implementation of ACC and DA provides conservative risk reductions. Second, there is no active control of the vehicle in the model. In many cases, an active ACC or DAS control would have initiated a speed reduction before the FCW or AEB was activated. Future studies should also investigate the effect of active control on simulations.

The fourth limitation is that context is not considered in either the collected data or the simulations. For the experiment, the presence of other road users (including lead vehicles) was not controlled for. That is, it may be that a driver had no other vehicles around when initiating a specific task, while the same driver when performing another task (or a different driver performing the same task) had other traffic around. This is likely to affect glance behaviours [47]. For the what-if simulations, there is a fundamental assumption that all of the tasks have an equal (and constant) probability of being initiated at all times except after looming (when thetadot exceeds 0.01 rad/s), when it is assumed that no off-road glances occur. Both of these context assumptions will affect the absolute estimates of risk, but they should have little effect on the main conclusions of the paper: that there is a need for the combined assessment of behaviour and ADASs. Future studies should consider context, accounting for the fact that there are different task initiation constraints for different tasks.

A fifth limitation is the assumption that all drivers will react directly after a threat is perceived (here operationalised as a brake reaction 0.5 s after thetadot reaches 0.01 rad/s). This assumption represents a simplistic driver response mode. One option for addressing this limitation is to simulate this condition as well, possibly by running the 6800 simulations of critical events with (for example) FCW + AEB + DA, but disable the driver brake response completely for a subset of the simulations. That is, disable the driver response to brake systems by 0–100% in steps of 10% – including disabling responses due to FCW, ACC, and DA (thus, only AEB being active). This would provide a conservative estimate of risk reductions of DA.

A sixth limitation is that the risk assessments here only include the rear-end crash scenario. Further studies should expand the application of the method to a variety of scenarios. To enable this, more research is needed to modelling driver response processes and glance behaviours in those other scenarios.

A seventh limitation is that the original crashes used in the simulations cannot, by definition, include crashes that are unique to the ADAS modality being studied. For example, DA systems might create some new crash type by their very existence—that was previously not there. For the rear-end scenario, it is hard to imagine such a crash type. However, if it exists, it cannot be analysed with our method.

An eighth limitation is that this study does not account for changes in driver behaviour when ADASs are used, other than the off-road glance behaviour. If, for example, an ADAS would prompt drivers to drive faster or keep a shorter headway (e.g. [48] shows reduced headway when driving with anti-lock brakes; ABS), the potential negative impact on the safety of such an adaptation is not covered in the current method. However, modifications to the seed kinematics could incorporate such aspects in the methodology. There are many other human behaviour changes (e.g. cognitive under load and overreliance) that may affect driving performance when ADASs are introduced. The impacts of such ADAS-induced driver behaviour changes are also currently not addressed in the method presented here. The research topic of overreliance on ADAS is currently large and ongoing regarding how best to capture the effects [49].

A ninth limitation is that the crash risks across tasks were compared without including the total task time in the calculations. That is, the exposure component of the risk increase (over baseline driving) for the total task duration was not taken into consideration when comparing tasks (see [25, 26] for details). It is possible to include total task-time-related exposure, but it was not included in this study as it would have complicated the presentation of the main conclusions, and would, in any case, have little effect on the conclusions of the study.

5 Conclusions

Corresponding to the two aims, this study (i) demonstrated the need for a holistic safety-impact assessment of the ADAS and (ii) implemented a substantial incremental development of the what-if (counterfactual) simulation methodology for carrying out such assessment. The results of the holistic safety impact of the combination of (i) the impact of the change in drivers' off-road glance behaviour due to the presence of the ADAS on, with (ii) the

IET Intell. Transp. Syst., 2020, Vol. 14 Iss. 9, pp. 1058-1067

1065

safety impact of the ADAS alone illustrates how the safety benefits of the ADAS may suppress the safety impact of changes in off-road glances that active assistance systems may produce, well below what is considered safe glance behaviour (e.g. Rockwell; [18]). That is, in the application of the method to a small dataset (for demonstration purposes), the changes to the glance distributions due to the presence of ADAS were very small in comparison with the overriding importance of the ADASs. As outlined in the limitations, the precision of these results should be seen as a first step approximation that will benefit from addressing the limitations. The results from the second set of simulations, which simulated the effect of varying degrees of ADAS failures, should be reasonably robust in its trends.

This study also demonstrates concerns with assessing glance behaviours in isolation, without considering the (potential) positive safety benefit of ADASs. The study has further demonstrated that the use of what-if (counterfactual) simulations can be one tool that can assess such combined effects – in a literature review, no other method to do this prospectively was found.

As a whole, this study is a further step towards a successively more accurate holistic risk assessment which includes driver behavioural responses such as glances and the safety effects of the ADAS. The use of what-if simulations has shown promise in previous scientific publications (e.g. [5, 24, 26, 27, 29–31, 50]) and this study provides further evidence of the relevance of the crash risk metric – demonstrating the usefulness of the methods in evaluating the combined risk of behaviour and system – a holistic risk assessment. As the method (and driver response models) evolves for more complex scenarios, it can be adapted to assess more complex systems (e.g. urban with cooperative systems) in the more complex scenarios.

With further evidence of real-life safety, for example, comparing simulation results with retrospective safety impact analyses of on-market systems, what-if simulations will be calibrated to real-life safety outcomes and become an effective safety tool for prospective assessment of behaviours and safety systems.

6 Acknowledgments

7 References

[1] Winner, H., Hakuli, S., Lotz, F., *et al.*: '*Handbook of driver assistance systems: basic information, components and systems for active safety and comfort*' (Springer, Cham, Switzerland, 2015)

[2] Benmimoun, M, Pütz, A., Zlocki, A., *et al.*: 'Impact assessment of adaptive cruise control (ACC) and forward collision warning (FCW) within a field operational test in Europe'. Presented at TRB, Washington, DC, January 2012

[3] Seppelt, B., Reimer, B., Angell, L., *et al.*: 'Considering the human across levels of automation: implications for reliance'. Proc. Ninth Int. Driv. Symp. Hum. Factors Driv. Assessment, Train. Veh. Des., Wermont, June 28th 2017

[4] Sayer, J., Bogard, S., Buonarosa, M.L., *et al.*: 'Integrated vehicle-based safety systems light-vehicle field operational test key findings report', The University of Michigan Transportation Research Institute, 2011

[5] Kusano, K.D., Gabler, H.C.: 'Safety benefits of forward collision warning, brake assist, and autonomous braking systems in rear-end collisions', *IEEE Trans. Intell. Transp. Syst.*, 2012, **13**, (4), pp. 1546–1555

[6] Cicchino, J.B.: 'Effectiveness of forward collision warning systems with and without autonomous emergency braking in reducing police-reported crash rates', Insurance institute for highway safety, 2016

[7] Brännström, M., Sjöberg, J., Coelingh, E.: 'A situation and threat assessment algorithm for a rear-end collision avoidance system'. Proceerings from 'Intelligent Vehicles Symp., 2008 IEEE', 2008 IEEE, Eindhovenm Netherlands, June 2008, pp. 102–107

[8] Fildes, B., Keall, M., Bos, N., *et al.*: 'Effectiveness of low speed autonomous emergency braking in real-world rear-end crashes', *Accid. Anal. Prev.*, 2015, **81**, pp. 24–29

[9] Ljung Aust, M., Jakobsson, L., Lindman, M.A., *et al.*: 'Collision avoidance systems - advancements and efficiency', SAE Technical Paper, 2015-01-1406, 2015

[10] Klauer, S.G., Guo, F., Simons-Morton, B.G., *et al.*: 'Distracted driving and risk of road crashes among novice and experienced drivers', *N. Engl. J. Med.*, 2014, **370**, (1), pp. 54–59

[11] Victor, T., Dozza, M., Bärgman, J., *et al.*: 'Analysis of naturalistic driving study data: SAFER glances, driver inattention, and crash risk', Transportation Research Board of the National Academy of Sciences, 2015

[12] Guo, F., Klauer, S.G., Fang, Y., *et al.*: 'The effects of age on crash risk associated with driver distraction', *Int. J. Epidemiol.*, 2017, **46**, (1), pp. 258–265

[13] Morando, A., Victor, T., Dozza, M.: 'Drivers anticipate lead-vehicle conflicts during automated longitudinal control: sensory cues capture driver attention and promote appropriate and timely responses', *Accid. Anal. Prev.*, 2016, **97**, pp. 206–219

[14] Morando, A., Victor, T., Dozza, M.: 'A reference model for driver attention in automation: glance behavior changes during lateral and longitudinal assistance', *IEEE Trans. Intell. Transp. Syst.*, 2018, pp. 1–11

[15] European_Communities: 'Commission recommendation of 26 may 2008 on safe and efficient in-vehicle information and communication systems; update of the european statement of principles on human machine interface', Commission of the European Communities, 2008

[16] NHTSA: 'Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices', National highway traffic safety administration at US Department of transportation, 2013

[17] Alliance_of_Automobile_Manufacturers: 'Statement of principles, criteria and verification procedures on driver interactions with advanced in-vehicle information and communication systems: including 2006 updated sections', Alliance of Automobile Manufacturers: Driver Focus-Telematics Working Group, 2006

[18] Rockwell, T.H.: 'Spare visual capacity in driving-revisited: new empirical results for an old idea', in Gale, A.G., Freeman, M.H., Haslegrave, C.M., *et al.* (Eds.): '*Second international conference on vision in vehicles*' (Elsevier, Nottingham, England, 1988), pp. 317–324

[19] Driver Focus-Telematics Working Group: 'Statement of principles, criteria and verification procedures on driver interactions with advanced In- vehicle information and communication systems', 2006

[20] Tivesten, E., Victor, T.: 'The timecourse of driver visual attention in naturalistic driving with adaptive cruise control and forward collision warning'. 4th Int. Driv. Distraction Ina. Conf., Sidney, New South Wales, November 2015

[21] Morando, A., Victor, T., Dozza, M.: 'A Bayesian reference model for visual time-sharing behaviour in manual and automated naturalistic driving', *IEEE Trans. Intell. Transp. Syst.*, 2019

[22] Reimer, B., Pettinato, A., Fridman, L., *et al.*: 'Behavioral impact of drivers' roles in automated driving'. Proc. of the 8th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications - Automotive'UI 16, Ann Arbor, MI, 2016

[23] Carsten, O., Lai, F.C.H., Barnard, Y., *et al.*: 'Control task substitution in semiautomated driving', *Hum. Factors*, 2012, **54**, (5), pp. 747–761

[24] Bärgman, J., Lisovskaja, V., Victor, T., *et al.*: 'How does glance behavior influence crash and injury risk? A 'what-if' counterfactual simulation using crashes and near-crashes from SHRP2', *Transp. Res. Part F Traffic Psychol. Behav.*, 2015, **35**, pp. 152–169

[25] Lee, J.Y., Lee, J.D., Bärgman, J., *et al.*: 'How safe is tuning a radio?: using the radio tuning task as a benchmark for distracted driving', *Accid. Anal. Prev.*, 2018, **110**, (Supplement C), pp. 29–37

[26] Bärgman, J., Boda, C.-N., Dozza, M.: 'Counterfactual simulations applied to SHRP2 crashes: the effect of driver behavior models on safety benefit estimations of intelligent safety systems', *Accid. Anal. Prev.*, 2017, **102**, pp. 165–180

[27] Lindman, M., Tivesten, E.: 'A method for estimating the benefit of autonomous braking systems using traffic accident data', SAE Tech. Pap., 2006

[28] Scanlon, J.M., Sherony, R., Gabler, H.: 'Using event data recorders from real-world crashes to investigate the earliest detection opportunity for an intersection advanced driver assistance system', SAE Technical Paper 2016-01-1457, Detroit, 2016

[29] Sander, U.: 'Opportunities and limitations for intersection collision intervention—A study of real world 'left turn across path' accidents', *Accid. Anal. Prev.*, 2017, **99**, pp. 342–355

[30] Isaksson-Hellman, I., Lindman, M.: 'The effect of a low-speed automatic brake system estimated from real life data', *Ann. Adv. Automot. Med./Annu. Sci. Conf.*, 2012, **56**, pp. 231–240

[31] Woodrooffe, J., Blower, D., Bao, S., *et al.*: 'Performance characterization and safety effectiveness estimates of forward collision avoidance and mitigation systems for medium/heavy commercial vehicles', The University of Michigan Transportation Research Institute for US Department of Transportation, 2012

[32] Viano, D.C., King, A.I., Melvin, J.W., *et al.*: 'Injury biomechanics research: an essential element in the prevention of trauma', *J. Biomech.*, 1989, **22**, (5), pp. 403–417

[33] Yang, K., Hu, J., White, N., *et al.*: 'Development of numerical models for injury biomechanics research: a review of 50 years of publications in the stapp car crash conference', *Stapp Car Crash J.*, 2006, **50**, pp. 429–490

[34] Wågström, L., Thomson, R., Pipkorn, B.: 'Structural adaptivity in frontal collisions: implications on crash pulse characteristics', *Int. J. Crashworthiness*, 2015, **10**, (4), pp. 371–378

[35] Palmonella, M., Friswell, M.I., Mottershead, J.E., *et al.*: 'Finite element models of spot welds in structural dynamics: review and updating', *Comput. Struct.*, 2005, **83**, (8–9), pp. 648–661

1066

IET Intell. Transp. Syst., 2020, Vol. 14 Iss. 9, pp. 1058-1067

[36] Brolin, K.: 'Human body models', in Tiwari, G., Mohan, D. (Eds.): '*Transport planning and traffic safety - making cities, roads, and vehicles safer*' (CRC Press, New York, NY, 2016), pp. 175–186

[37] CARTRE-consortium: 'Position paper on human factors', 2018

[38] Bellet, T., Bornard, J.-C., Bertrand, R.*, et al.*: 'Use of a cognitive simulation model of the driver to support the Virtual Human Centred Design (V-HCD) of ADAS and automated vehicles', Graz Symposium Virtual Vehicle: Artifical Intelligence Meets Model-Centric Design

[39] Blatt, A., Pierowicz, J.A., Flanigan, M.*, et al.*: 'SHRP2 safety research - Naturalistic driving study: Field data collection', Transportation Research Board at the National Academy of Sciences, 2015

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

[41] Bradley, M.M., Lang, P.J.: 'Measuring emotion: the self-assessment manikin and the semantic differential', *J. Behav. Ther. Exp. Psychiatry*, 1994, **25**, (1), pp. 49–59

[42] Markkula, G., Engström, J., Lodin, J.*, et al.*: 'A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies', *Accid. Anal. Prev.*, 2016, **95**, (Part A), pp. 209–226

[43] Engström, J., Bärgman, J., Nilsson, D.*, et al.*: 'Great expectations: a predictive processing account of automobile driving', *Theor. Issues Ergon. Sci.*, 2018, **19**, (2), pp. 156–194

[44] Clark, A.: '*Surfing uncertainty: prediction, action, and the embodied mind*' (Oxford University Press, New York, NY, 2016)

[45] JAMA.: 'Guideline for In-vehicle display systems - version 3.0', Jama, 2004

[46] Neale, V.L., Dingus, T.A., Klauer, S.G.*, et al.*: 'An overview of the 100-car naturalistic study and findings', 19th Int. Technical Conf. on the Enhanced Safety of Vehicles (ESV); Paper nr. 05-0400, NHTSA, 2005

[47] Tivesten, E., Dozza, M.: 'Driving context influences drivers' decision to engage in visual–manual phone tasks: evidence from a naturalistic driving study', *J. Safety Res.*, 2015, **53**, pp. 87–96

[48] Sagberg, F., Fosser, S., Saetermo, I.-A.: 'An investigation of behavioural adaptation to airbags and antilock brakes among taxi drivers', *Accid. Anal. Prev.*, 1997, **29**, (3), pp. 293–302

[49] Merat, N., Sepplet, B., Louw, T.*, et al.*: 'The 'out-of-the-loop' concept in automated driving: proposed definition, measures and implications', *Cogn. Technol. Work*, 2018, **21**, (1), pp. 87–98

[50] Page, Y., Fahrenkrog, F., Fiorentino, A.*, et al.*: 'A comprehensive and harmonized method for assessing the effectiveness of advanced driver assistance systems by virtual simulation: the P.E.A.R.S. Initiative'. 24th Int. Technical Conf. on the Enhanced Safety of Vehicles (ESV); Paper nr. 15–0370' US National Higway Traffic Safety Administration, 2015

IET Intell. Transp. Syst., 2020, Vol. 14 Iss. 9, pp. 1058-1067

1067