Accident Analysis and Prevention 97 (2016) 162–175

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Cooperative warning systems: The impact of false and unnecessary alarms on drivers' compliance

Frederik Naujoks[a,*], Andrea Kiesel[b], Alexandra Neukum[a]

[a] University of Wuerzburg, Center for Traffic Sciences (IZVW), Germany
[b] University of Freiburg, Department of Psychology, Germany

ARTICLE INFO

Article history:
Received 19 April 2016
Received in revised form 6 September 2016
Accepted 9 September 2016
Available online 14 September 2016

Keywords:
Advisory warnings
False alarms
Cry wolf effect
Warning urgency

ABSTRACT

Cooperative warning systems have a great potential to prevent traffic accidents. However, because of their predictive nature, they might also go along with an increased frequency of incorrect alarms that could limit their effectiveness. To better understand the consequences associated with incorrect alarms, a driving simulator study with $N = 80$ drivers was conducted to investigate how situational context and warning urgency jointly influence drivers' compliance with an unreliable advisory warning system (AWS). The participants encountered several critical urban driving situations and were either assisted by a 100% reliable AWS, a 60% reliable AWS that generated false alarms (without obvious reason) or a 60% reliable AWS that generated unnecessary alarms (with plausible reason). A baseline drive without any assistance was also introduced to the study. The warnings were presented either only visually or visual-auditory. In line with previous research, drivers' compliance and effectiveness of the AWS was reduced by false alarms but not by unnecessary alarms. However, this so-called *cry wolf effect* (Breznitz, 1984) was only found in the visual-auditory condition, whereas there was no effect of warning reliability in the condition with visual AWS. Furthermore, false but not unnecessary alarms caused the participants to rate the AWS less favourably during a follow-up interview. In spite of these negative effects of false alarms, a reduction in the frequency of safety-critical events (SCEs) and an earlier braking onset were evident in all assisted drives compared with that of non-assisted driving, even when the AWS was unreliable. The results may thus lower concerns about the negative consequences of warning drivers unnecessarily about upcoming traffic conflicts if the reasons of these alarms are comprehensible. From a perspective of designing AWS, we recommend to use less urgent warnings to prevent the *cry wolf effect*.

1. Introduction

1.1. Background and scope of the study

In Germany, most accidents with personal injury are reported in urban areas with most of them happening in complex driving situations, such as intersections (Federal Statistical Office, 2013). Consequently, assisting drivers in these situations has a great potential to improve traffic safety. However, driver assistance that is merely based on on-board perception of the vehicle environment, such as a camera or radar, may not be capable of sufficiently analysing the driving environment in these conflict situations to provide comprehensive driver support (Seeliger et al., 2014). Specifically, the recognition and tracking of vulnerable road users and that of partly or even fully occluded road users is necessary to assist the driver efficiently (Edquist et al., 2012; Hamdar et al., 2016; Marciano and Yeshurun, 2015; Naujoks et al., 2015a,b; Rogé et al., 2012).

Driver assistance based on cooperative perception (e.g., car-to-car or car-to-infrastructure communication) has received considerable interest because it may provide a solution to these technical limitations by fusing vehicle localised environmental perception with information provided by other road users or the infrastructure (e.g., sensors mounted to traffic lights at intersections). Several national (e.g., sim^{TD} and Ko-FAS) and European research projects (e.g., INTERSAFE and DRIVE C2X) have thus dealt with the technological advancement of *cooperative perception*. Similar technological developments, like research on so-called *connected vehicles*, take place on an international level. For example, the US Department of Transportation currently runs a Connected

* Corresponding author.
E-mail addresses: naujoks@wivw.de, naujoks@psychologie.uni-wuerzburg.de (F. Naujoks).

http://dx.doi.org/10.1016/j.aap.2016.09.009

Vehicle Deployment Program on three deployment sites (USDOT, 2015).

A promising application of cooperative perception is to inform drivers about impending conflict situations, even if the conflicting road user is occluded from the host vehicle's point of view (i.e., *advisory warnings*, see Seeliger et al., 2014; Naujoks et al., 2015a,b; Maag et al., 2015). The potential of these predictive warning systems to enhance driving safety has been shown repeatedly in different contexts, such as congestion tail warnings (Totzke et al., 2012; Werneke et al., 2013; Winkler et al., 2016), obstacle warnings (Mahr et al., 2010) and intersection collision warnings (Lenné et al., 2008; Naujoks and Neukum, 2014a; Zarife, 2014).

So far, these studies have dealt with perfectly reliable warning systems. However, assisting the driver by means of cooperative perception technology requires the modelling and prediction of the traffic situation based on fused sensor data. Due to the probabilistic nature of this cooperative perception approach, cooperative warning systems will not be perfectly reliable, which might in turn decrease the warnings' usefulness. In a literature review, Wickens and Dixon (2007) concluded that warning systems that are less than 70% reliable (with a confidence limit of roughly 14%), were not beneficial for human operators.

Reliability of warning systems can influence driver behaviour in several ways. Previous research has repeatedly revealed that false alarms reduce compliance with warning systems, for example by increasing reaction times, a phenomenon that has been labelled the *cry wolf effect* (Breznitz, 1984; Getty et al., 1995; Sorkin, 1988). The *cry wolf effect* has been named after the fable by Aesop about the "The boy who cried wolf"; however, in human factors psychology, it is not a young shepherd boy that has warned of a wolf worrying the sheep too often but an automated warning system that has issued too many false alarms (Roulston and Smith, 2004). The reason for a decreased compliance caused by false alarms may be a diminished contingency between the presentation of a stimulus, i.e., the driver warning, and the need for a reaction (Kiesel and Miller, 2007). The *cry wolf effect* is well-known from other applied research settings, such as aviation (Pritchett, 2001), medicine (Kestin et al., 1988; Meredith and Edworthy, 1995) or process management (Kragt and Bonten, 1983; Lee and Moray, 1992). In the worst case, users ignore or even switch off warning systems because the rate of false alarms is too high. In contrast to false alarms, missed alarms represent situations in which a traffic conflict that the driver should attend to does not trigger a warning signal. Missed alarms can lead to increased reaction times to critical situations as well as an increased crash risk compared with non-assisted driving (Mahr et al., 2010; Yamada and Kuchar, 2006).

The present paper revisits the issue of how false alarms influence cooperative warning effectiveness. By means of cooperative perception, drivers are assisted by an *advisory warning system* (AWS) in situations in which the conflicting road users are occluded when the driver approaches (e.g., by parked cars on the side of the road); this situation can only be resolved by cooperative warning systems. However, precisely in these situations, the reliability of the warning system might largely influence drivers' compliance because drivers cannot directly verify the correctness of the alarm. Within this context, two research questions are investigated that have not been comprehensively dealt with by prior research, namely, the presence of situational indicators from which the driver can infer the reason why a false alarm has occurred and the impact of the warning urgency.

1.2. False alarms effects on warning effectiveness

To date, most studies dealing with the *cry wolf effect* in the context of driver warning systems have focused on situations in which false alarms are presented without any reason (e.g., Bliss and Acton, 2003; Cummings et al., 2007; Yamada and Kuchar, 2006). These types of alarms may be most likely due to technical failures of the warning system. Alternatively, it is also possible that an alarm is issued because the warning system detected a potential conflict situation; however, as the situation develops further, the conflict resolves such that there is no imminent accident risk from the driver's point of view. These types of unnecessary alarms, thus, represent failures of the situational analysis and prediction involved in the generation of the warning (Weidl and Breuel, 2012; Weidl et al., 2013). For example, a pedestrian that is standing on the side of the road and is about to cross into the road may trigger an alarm, but the pedestrian may finally not enter the road as the host vehicle approaches. Consequently, the driver may not have to react to the warning at all.

According to Lees and Lee (2007), unnecessary alarms differ from "truly" false alarms because they provide the opportunity to understand the *process* of the warning system (cf. Lee and See, 2004). Understanding the *process* involved in the generation of the warning may increase trust in the warning system (Lee and See, 2004) and may consequently prevent diminished compliance despite unnecessary alarms (Lees and Lee, 2007). The distinction between false alarms and unnecessary alarms is illustrated in Table 1.

One may assume that the *cry wolf effect* would not occur in situations with unnecessary alarms because the driver is able to understand the situational determinants of unnecessary alarms (cf. Dzindolet et al., 2003). For example, in a study by Cotté et al. (2001), alarms of a collision warning system warned drivers of obstacles during a simulated drive. Thereby, false alarms occurred either without any obvious reason or unnecessary alarms occurred in situations in which objects on the side of the road could have caused the alarms. False and unnecessary alarms both decreased compliance with the warning system. These results contradict the results of Lees and Lee (2007) who reported that only false alarms, not unnecessary alarms, had a detrimental effect on drivers' compliance. These mixed empirical results beg the question whether false and unnecessary alarms cause different consequences on drivers' compliance and subjective evaluation of the warning system, which will be revisited in this study. In addition, we explore the impact of warning urgency on false and unnecessary alarms because we hypothesize that more urgent but false alarms are more detrimental for drivers' compliance.

1.3. Urgency and false alarm effects

From the driver's point of view, using more than one modality in the alarm design delivers a higher level of subjective urgency (Politis et al., 2013). Accordingly, *imminent collision warnings* in automobiles are usually presented in more than one modality to decrease reaction times and to enhance alarm effectiveness (ISO 15623:2013). Multimodal presentation of warning signals can speed up the cognitive processes involved in the selection and execution of an appropriate response, such as braking or steering. The advantageous effect of presenting more than one stimulus at once that requires a reaction, so-called *redundancy gain*, has been demonstrated repeatedly in cognitive psychology research (cf., Miller, 1982; Raab, 1962). Another goal of multimodal warning systems is to draw the driver's attention to a visual display on which relevant information is presented if the driver's gaze is not oriented towards that direction (Campbell et al., 2007). Consequently, the advantage of a multimodal warning system in the context of *imminent collision warnings* has been demonstrated in various studies (e.g., Ho et al., 2007; Kramer et al., 2007).

Nevertheless, warning urgency may be an important design aspect of cooperative warning systems. The previously mentioned studies mainly focussed on *imminent collision warnings* (Lenné and

Table 1
Illustration of the distinction between false alarms, unnecessary alarms and correct alarms (according to Lees and Lee, 2007).

False alarm: No pedestrian present	Unnecessary alarm: Pedestrian is present but does not cross	Correct Alarm: Crossing pedestrian
		

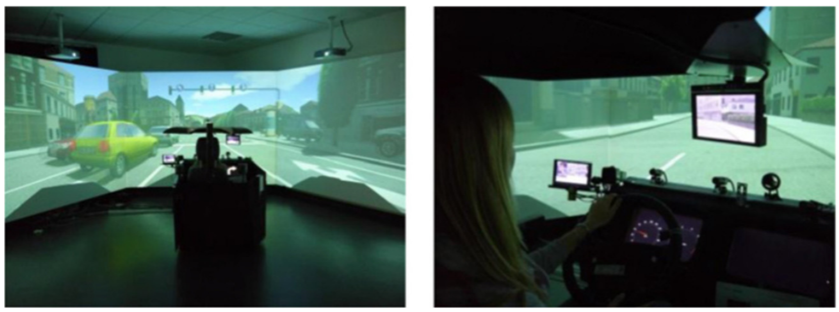


Fig. 1. Static driving simulator of the WIVW.

Triggs, 2009) that were usually presented just within the driver's reaction time but not earlier (ISO 15623:2013, cf. Maag et al., 2015), with the goal to induce an immediate driver reaction. The cooperative perception approach enables the presentation of *advisory warnings* (Lenné and Triggs, 2009) at an earlier stage of the conflict situation (Seeliger et al., 2014; Weidl et al., 2013) with the aim to direct the driver's attention and render him/her ready to respond (Naujoks et al., 2015a,b). However, these alarms do not require an immediate reaction. It is thus questionable whether using urgent alarms, as in case of *imminent crash warnings*, is useful in this case, especially as false alarms may be more annoying when presented with high urgency.

It has been shown that the *cry wolf effect* is diminished by using graded warning systems instead of binary warning systems (Gupta et al., 2002; Lee et al., 2004). In a wider context, this alarm design has been labelled as a *likelihood alarm system* (Sorkin et al., 1988; Sorkin, 1988). According to this approach, the urgency of the warning signal corresponds to its confidence level, for example by using warning tones of different loudness or different warning symbols (Finger and Bisantz, 2002), making it easier to distinguish between alarms that are likely to require a response and those that are unlikely to require a response (Woods, 1995; Montgomery and Sorkin, 1996; Montgomery, 1999). This may help to maintain the driver's readiness to respond even if the warning system is not perfectly reliable. The *cry wolf effect* may thus be prevented if the warning is not presented as urgent as in a case of *imminent collision warnings*. Accordingly, Hoffmann and Gayko (2012) have argued that urgent driver warnings, as opposed to more informative warnings, are less "pardonable" from a driver's perspective. Thus, the impact of false alarms on drivers' compliance will be investigated in this study with two different versions of AWS that vary in urgency.

1.4. Research question and hypothesis

The focal point of this study was the investigation of the interplay between the urgency of *advisory warnings* based on cooperative perception and the impact of different levels of warning reliability. We expected that the effectiveness of the warning system would be impaired in case of "truly" *false alarms* that were presented without any obvious reason but not as much in a case of *unnecessary alarms* that were presented in the presence of potential conflict situations that the driver did not have to react to. For this, we compared a 100% reliable AWS, a 60% reliable AWS that generated false alarms and a 60% reliable AWS that generated unnecessary alarms. In addition, we investigated whether presenting advisory warnings only in the visual modality or, in a more imminent visual-auditory way, would change the way that drivers responded to the warning system's unreliability. The latter question has only been dealt within the context of *missed alarms* (i.e., no warning presentation although a collision is imminent; Shah et al., 2015) and not in the context of *false* and *unnecessary alarms*. To assess whether drivers would still benefit from the AWS even if it is not perfectly reliable, a non-assisted baseline group was also part of the study.

During the study, the participants encountered four different traffic scenarios (see the method section) that resulted in critical situations if the drivers did not react by braking. To assess drivers' compliance with the AWS, we followed Meyer (2004, p. 199) who defined compliance with a warning as "*. . .the response when an operator acts according to a warning signal and takes some evasive action. . .*". Consequently, compliance with the warnings was measured by Time-to-arrival (TTA, see Method Section 2.6 and Fig. 2) at braking onset and the maximum brake pedal position. Warning effectiveness was investigated via the frequency of situations with a minimum TTA of below one second, as these situations can be

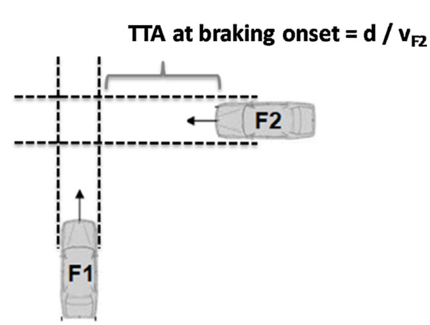


Fig. 2. Calculation of the TTA at braking onset. The value was calculated as soon as the participant initiates the braking.

classified as *near misses* (Hayward, 1972). In addition, participants' subjective evaluation of the warning system's *usefulness*, *ease of use*, *trust* and *acceptance* were compared between the experimental groups.

2. Method

2.1. Setting

Participants completed a simulator drive during which they encountered various traffic conflicts while being assisted by different variants of a cooperative AWS or while driving without any assistance. The study was performed in a static driving simulator at the Wuerzburg Institute for Traffic Sciences (WIVW, Fig. 1). The driving simulator provided a 300° horizontal field of view via five image channels. Two LCD displays were used to show the rear-view and the left outside mirror. The vehicle mock-up included a force feedback steering wheel.

2.2. Human-machine interface (HMI) and driving situations

The design of the AWS was based on prior research. During the approach to the traffic conflicts, the advisory warnings were presented two seconds before the drivers had to initiate hard braking, assuming a constant deceleration of $-8\,\text{m/s}^2$ (Naujoks and Neukum, 2014a). Warning timing was based on prior studies that showed this timing was feasible from an analytical (Weidl et al., 2013), practical (Seeliger et al., 2014) and human factors perspective (Naujoks and Neukum, 2014a). Warnings were presented visually on a simulated Head-Up Display (HUD). In the visual-auditory conditions, warnings were accompanied by a non-obtrusive tone (500 Hz sinus; for a similar procedure, see Naujoks et al., 2015a,b). The visual element pictured the traffic situation and the direction from which the conflicting road user was approaching. The usefulness of this spatial warning approach has also been evaluated in an earlier study (Naujoks and Neukum, 2014b) and was in line with a growing body of literature that investigated spatial driver warnings (e.g., Zarife, 2014; Weller et al., 2013).

Four different driving situations were used (see Table 2) and they were taken from prior studies that had already investigated use cases of cooperative warning systems (Naujoks et al., 2015a; Naujoks and Neukum, 2014a). While driving on the designated route, the drivers had to react to a turning vehicle taking the drivers' right of way (situation "turning vehicle"), to a crossing cyclist taking the drivers' right of way (situation "crossing cyclist"), to a crossing vehicle taking the drivers' right of way (situation "crossing vehicle") and to a pedestrian entering the road between parked cars (situation "crossing pedestrian"). All situations were timed so that participants would collide with the conflicting road user if their travel speed was not reduced. During the approach to all situations, the drivers' sight of the conflicting road users was occluded so that anticipatory driving reactions were hardly possible (cf. Naujoks et al., 2015a,b). The visual HUD element of the HMI is also presented in Table 2 next to the corresponding traffic situation.

2.3. Experimental design and trial composition

A mixed $3 \times 2 \times 4$ between-within-subjects design was used with the between-subjects factors *warning reliability* (three levels) and *warning urgency* (two levels) and the within-subject factor *driving situation* (four levels). The AWS was either perfectly reliable, in the sense that advisory warnings were only given in situations that required a reaction by the driver or the frequency of correct warnings was reduced to 60%. Missed alarms were not part of the study design (i.e., all critical situations were detected correctly, even in the conditions with reduced reliability). We expected the reliability level to meet boundary conditions in which the frequency of incorrect alarms would influence compliance, but would not render the AWS unusable. Wickens and Dixon (2007) concluded that 70% reliability (with a 95% confidence limit of about 14%) are needed for warning systems to be beneficial for human operators, regardless of whether the reliability is reduced by false or missed alarms.

For the AWS with reduced reliability, the drivers experienced system failures either in situations without any obvious reason (factor level: 60% reliable with false alarms) or with another road user being present that could have caused the warning (factor level: 60% reliable with unnecessary alarms). Advisory warnings were either presented only visually or visual-auditory. In addition to these experimental groups, a control condition was introduced in which participants completed the simulator drive without any assistance (Baseline drive). Regarding the within-subject factor driving situation, all drivers completed the four driving situations.

The simulator drive included each of the four critical situation three times; thus, each participant encountered a total number of twelve conflict situations. For the AWS conditions with reduced reliability, the drivers additionally experienced eight system failures during the drive, resulting in a reliability of 60% (i.e., 12 correct alarms, 8 incorrect alarms). In the condition with false alarms, system failures consisted of the AWS warning without any obvious reason. In the condition with unnecessary alarms, another road user was present that could have caused the warning, such as a pedestrian standing on the side of the road or a cyclist waiting at an intersection (see Table 1). Drivers in the condition with reliable AWS and those in the Baseline condition also completed eight similar road segments (without warnings) to keep the length of the test drive comparable between the drivers. To reduce the possibility of anticipating the traffic conflicts, we added twelve non-critical situations that did not trigger any warning and resembled the test situations on a superficial level (e.g., same road geometry and same buildings on the side of the road) into the simulator run.

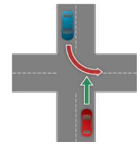
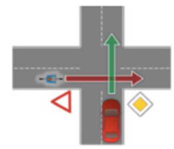
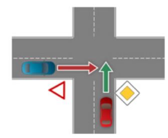
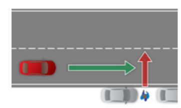
Learning effects may influence compliance with warning systems in several ways. First, drivers might behave differently when they are not familiar with a warning system (Aust et al., 2013). Therefore, the drivers completed the four critical driving situations once with the help of a perfectly reliable AWS during a separate training drive. Second, learning effects may influence how drivers adapt their behaviour to incorrect warnings. For example, compliance with the AWS may be reduced especially after the first incorrect warning. However, subsequently presented correct warnings may counteract negative effects on compliance. It is also possible that the time passed between an incorrect alarm and a subsequent correct alarm may have an effect on compliance. Therefore, the order of the test situations (12 critical situations, 8 false/unnecessary warnings and 12 non-critical situations) in the main part of the study was randomised with the constraint that the last situation was always a critical one. In this way, learning effects were assumed to affect the drivers' reactions to the correct warnings equally.

2.4. Sample

A total of $N = 81$ drivers (43 female) were recruited for participation in the study. One participant could not finish the study because of simulator sickness, resulting in a sample of $N = 80$. The drivers were assigned randomly to the experimental conditions with a cell size of $n = 10$ drivers, while $n = 20$ drivers were assigned to the baseline group (see Table 3).

The participants were selected from the WIVW test driver panel and had received extensive simulator training prior to the start of the study. During this training procedure, all participants completed two consecutive training sessions, each lasting approximately 2 h, in which they were trained to handle the simulated

Table 2
Overview of driving situations and corresponding visual warning elements presented in the HUD. The host vehicle is depicted in red and the conflicting road user is depicted in blue.

Situation and description	Overview	HUD
Turning vehicle: An oncoming vehicle takes the driver's right of way at an intersection. The driver is going straight and the conflicting vehicle is turning in. The sight of the turning vehicle is occluded by oncoming traffic.		
Crossing cyclist: A cyclist coming from the left is taking the driver's right of way. The driver is going straight. The sight of the crossing cyclist is occluded by parked cars on the side of the road.		
Crossing vehicle: A vehicle coming from the left is taking the driver's right of way. The driver is going straight. The sight of the crossing vehicle is occluded by parked cars on the side of the road.		
Crossing pedestrian: A pedestrian is entering the road between parked cars and crosses it. The sight of the pedestrian is occluded by parked cars on the side of the road.		

vehicle correctly. A description of the training can be found in Hoffmann et al. (2003) or Naujoks et al. (2015a,b). The mean age of the drivers was 38 years (*SD* = 15, *MIN* = 20, *MAX* = 71, lower quartile = 25, upper quartile = 53) with a mean driving experience of 19 years (*SD* = 14, *MIN* = 2, *MAX* = 50, lower quartile = 6, upper quartile = 30). On average, they had a mileage of 14,200 km in the preceding year (*SD* = 12200, *MIN* = 300, *MAX* = 70,000; lower quartile = 6300, upper quartile = 20,000). While the sample was quite heterogeneous with regard to these sample characteristics, special groups of drivers such as novice drivers or drivers with very low mileage were only single cases. Using one and a half of the quartile range to search for outliers (Tukey, 1977), three outliers on the mileage variable were identified (one case with 50,000 km in the baseline, one case with 50,000 km in the visual condition with 60% false alarms and one case with 70,000 km in the visual-auditory condition with 60% unnecessary alarms). The experimental groups did not differ significantly in any of the sample characteristics (age: $F(6,73) = 0.31$, $p = 0.931$; driving experience: $F(6,73) = 0.28$, $p = 0.946$; mileage: $F(6,73) = 1.08$, $p = 0.380$; between-subjects ANOVAs comparing the experimental groups were conducted).

2.5. Procedure and instructions

Upon arrival at the test site, the drivers were familiarised with the AWS prior to the main part of the study. The drivers in the experimental conditions were informed that they would be assisted by a new type of driver assistance system, but they were not given any precise information about the AWS (such as interpretation of the HMI, information about the reliability or about the possibility of false alarms). Afterwards, a training drive was completed that contained the four critical situations. The drivers in the experimental conditions were assisted by a perfectly reliable warning system in this drive. Advisory warnings were presented either visually or visual-auditory, depending on the experimental condition. Drivers in the control condition were not assisted by the AWS in this part of the study (i.e., they completed the four critical situations without any assistance).

After system familiarisation, the main part of the study was conducted. The drivers were instructed to complete the simulator course in a timely manner and to follow traffic rules. No instruction about the reliability of the AWS was given to the drivers. Completing the drive took the participants approximately 45 min, regardless of the experimental condition. The situations were completed in randomised order. After the test drive, drivers completed a follow-up questionnaire. The participants received financial compensation of 50 euros.

Table 3
Between-subjects factors and cell size per condition.

Urgency	Reliability of AWS		
	100%	60% with false alarms (FA)	60% with unnecessary alarms (UA)
Visual-auditory	10	10	10
Visual	10	10	10
Baseline (no AWS)	20		

2.6. Dependent measures

Compliance with the AWS was measured by the Time-to-arrival (TTA) at braking onset and the maximum brake pedal position. TTA at braking onset represents the time that remains to avoid the collision given that the host vehicle continues on its path without any velocity change. Greater values, thus, signify that the drivers initiated the braking process with a larger time budget and complied more strongly with the AWS. It is calculated by dividing the host vehicle's distance to the conflict point with the conflicting road user (e.g., vehicle crossing the intersection), by the current velocity of the host vehicle at the moment of braking onset. To describe the drivers' reactions more fully, the maximum brake pedal position was also taken into account, as it may have been possible that the

Table 4
Rating items contained in the follow-up questionnaire.

Dimension	Item number	Item
Ease of use (Venkatesh et al., 2003)	1	Reading the warnings was strenuous.
	2	The warnings were annoying.
	3	The warnings were disturbing.
	4	The warnings were unnecessary.
	5	The warnings distracted me from driving.
	6	The warnings were understandable.
Usefulness (Venkatesh et al., 2003)	7	Because of the warnings I could recognize traffic conflicts earlier.
	8	The warnings were helpful.
	9	Driving was safer because of the warning system.
	10	The warning system will reduce traffic accidents.
Trust in automation (Jian et al., 2000)	11	Because of the warning I endangered other road users.
	12	The warning system was reliable.
Intention to use (Venkatesh et al., 2003)	13	I would use the warning system during every day driving.

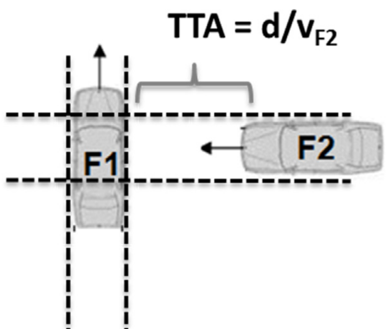


Fig. 3. Calculation of the minimal TTA. The value was only calculated when the respective conflicting road user is in the path of the host vehicle.

drivers compensated a later braking onset with a stronger brake reaction.

Warning effectiveness was measured by the frequency of safety critical events (SCEs) during the test situations. Here, it was assessed by whether a minimum TTA of less than one second to the conflicting road user was registered during the event, as this threshold has been used to define critical driving events or *near misses* (Hayward, 1972) in prior research (Hayward, 1972; van der Horst, 1990; Naujoks and Totzke, 2014).[1]

A follow-up questionnaire contained several rating items that aimed to assess the driver's subjective evaluation of the warning system (see Table 4). *Ease of use*, *usefulness*, *trust* and *intention to use* (see, e.g., Davis, 1989; Gold et al., 2015; Venkatesh et al., 2003) the warning system were measured by several items that were specifically constructed for use in the study. The drivers were asked to indicate how much they would agree to the statements shown in Table 4 on a five-point scale ranging from "not at all" to "very much". They were also given the opportunity to state that they "don't know" whether they agreed to the statement in an extra category.

[1] An alternative measurement would have consisted of comparing the mean TTA_{min}-values between the experimental conditions without converting it into a binary measure. This approach was, however, not feasible, as it leads to outliers especially during assisted driving. Specifically, in the case that the driver reacts rapidly to the warning and reduces the velocity considerably before the conflicting road user is entering her/his path (see Fig. 3), very large TTA_{min}-values are produced that cannot be meaningfully used to compare average TTA_{min}-values between the experimental conditions.

3. Results

3.1. Missing values

In one of the situations, one participant braked too early in the approach so that the test situation could not be produced (situation: "crossing pedestrian", experimental condition: visual-auditory with false alarms). In the follow-up interview, one driver stated that he/she did not know whether "The warning system will reduce traffic accidents".

3.2. Compliance

We were primarily interested in how false and unnecessary alarms would influence the participants' reactions to correct alarms during the study. Thus, we compared compliance to the twelve correct alarms between the experimental groups. The compliance measures collected in the study, i.e., TTA at braking onset and maximum brake pedal position, were subject to a full-factorial mixed $3 \times 2 \times 4$ between-within ANOVA with the between-subjects factors *warning reliability* (100% vs. 60% with false alarms vs. 60% with unnecessary alarms) and *warning urgency* (visual vs. visual-auditory) and the within-subject factor *driving situation*. The values were additionally compared with the control group that completed the situations without any warning using planned contrasts.

In a first step, the frequency of trials in which the drivers actually responded to the situations by braking was analysed because a decreased frequency of brake reactions could be interpreted as lowered compliance with the AWS. Table 5 shows that most drivers reacted by braking in the assisted drives, with only minor variations between the reliability and urgency conditions. In most cases, the drivers braked in all the three repetitions of the respective scenarios, as shown in the last column of the table. For the further analyses, the values of TTA at braking onset and maximum brake pedal position were averaged per situation and per participant (i.e., the values from the repeated exposure to the situations were averaged for each participant) to keep the cell size for the mixed between-within ANOVAs on a comparable level and to prevent drop-out due to incomplete data. Only one driver had to be excluded from the analysis because he/she did not show a brake reaction in all three repetitions of the scenario "turning vehicle" (experimental condition: visual-auditory, 100% reliable).

The braking onset is depicted in Fig. 4. Regarding the influence of false and unnecessary alarms, there was both a marginally significant main effect of *warning reliability* and a marginally significant interaction between *warning reliability* and *warning urgency* (see Table 6). Separate ANOVAs for each urgency condition showed that false alarms led to a later braking onset in the visual-auditory

Table 5
Frequency of valid trials in which drivers reacted by braking; FA = false alarms, UA = unnecessary alarms.

Situation	Warning urgency	Warningreliability	N	Trials	Trials with brake reaction	%	Number of brake reactions/driver			
							0	1	2	3
Turning vehicle	Baseline	–	20	60	47	78	0	4	5	11
	visual-auditory	100%	10	30	27	90	1	0	0	9
	visual-auditory	60% with FA	10	30	29	97	0	0	1	9
	visual-auditory	60% with UA	10	30	29	97	0	0	1	9
	visual	100%	10	30	30	100	0	0	0	10
	visual	60% with FA	10	30	29	97	0	0	1	9
	visual	60% with UA	10	30	30	100	0	0	0	10
Crossing vehicle	Baseline	–	20	60	53	88	0	2	3	15
	visual-auditory	100%	10	30	28	93	0	1	0	9
	visual-auditory	60% with FA	10	30	29	97	0	0	1	9
	visual-auditory	60% with UA	10	30	28	93	0	0	2	8
	visual	100%	10	30	30	100	0	0	0	10
	visual	60% with FA	10	30	29	97	0	0	1	9
	visual	60% with UA	10	30	30	100	0	0	0	10
Crossing pedestrian	Baseline	–	20	60	60	100	0	0	0	20
	visual-auditory	100%	10	30	29	97	0	0	1	9
	visual-auditory	60% with FA	10	30	29	97	0	0	1	9
	visual-auditory	60% with UA	10	30	29	97	0	0	1	9
	visual	100%	10	30	30	100	0	0	0	10
	visual	60% with FA	10	30	30	100	0	0	0	10
	visual	60% with UA	10	30	30	100	0	0	0	10
Crossing cyclist	Baseline	–	20	60	58	97	0	0	2	18
	visual-auditory	100%	10	30	30	100	0	0	0	10
	visual-auditory	60% with FA	10	30	30	100	0	0	0	10
	visual-auditory	60% with UA	10	30	30	100	0	0	0	10
	visual	100%	10	30	30	100	0	0	0	10
	visual	60% with FA	10	30	30	100	0	0	0	10
	visual	60% with UA	10	30	30	100	0	0	0	10

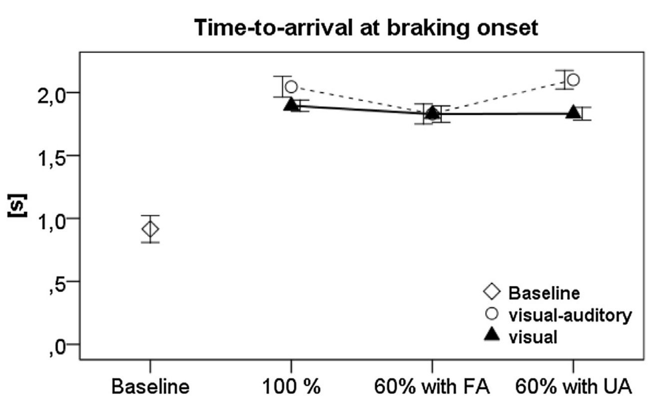


Fig. 4. TTA at braking onset, mean and 95%-confidence limits are depicted; FA = false alarms, UA = unnecessary alarms.

condition (Fig. 4, effect *warning reliability*: $F(2,26) = 4.21$, $p = 0.026$, $\eta^2 = 0.24$, difference to average braking onset: $M = -0.13$ s, $SE = 0.06$, $p = 0.022$), while unnecessary alarms even led to a faster braking onset (difference to average braking onset: $M = 0.14$ s, $SE = 0.06$, $p = 0.017$). There was no influence of *warning reliability* when the advisory warnings were presented only visually ($F(2,27) = 0.63$, $p = 0.543$, $\eta^2 = 0.04$). It was also evident that the participants began the braking process at a larger distance to the conflict point when the advisory warnings were presented visual-auditory ($M = 1.96$ s, $SE = 0.03$) than when they were presented only visually ($M = 1.85$ s, $SE = 0.03$, main effect *warning urgency*); however, Fig. 4 shows there was no such difference between the reliability levels with false alarms. The effects were independent of the driving *situation*, as there was no main effect or interaction effect with the factor *situation*. As can be observed from Fig. 4 and Table 7, drivers started the braking process earlier when they were assisted by the AWS, independent of the reliability and urgency of the warnings.

The average maximum brake pedal positions are depicted in Fig. 5. Higher maximum brake pedal positions were observed with visual warnings ($M = 46.42\%$, $SE = 3.13$) than with visual-auditory warnings ($M = 32.60\%$, $SE = 3.19$, main effect *warning urgency*, see Table 8). The maximum brake force applied by the drivers varied between the situations (main effect *situation*), with the strongest brake reactions being observed in the situation "crossing pedestrian" ($M = 47.41\%$, $SE = 2.85$; situation "turning vehicle": $M = 38.01\%$, $SE = 2.54$; situation "crossing vehicle": $M = 35.76\%$, $SE = 2.23$; situation "crossing cyclist": $M = 36.86\%$, $SE = 2.24$). There was no main effect of *warning reliability* on the

Table 6
TTA at braking onset, ANOVA results.

Effect	F	df_1	df_2	p	η^2
Warning urgency	5.23	1	53	0.026	0.09
Warning reliability	3.03	2	53	0.057	0.10
Warning reliability * Warning urgency	3.03	2	53	0.057	0.10
Situation	28.77	3	51	<0.001	0.63
Situation * Warning reliability	0.98	6	104	0.443	0.05
Situation * Warning urgency	1.03	3	51	0.389	0.06
Situation * Warning reliability * Warning urgency	1.11	6	104	0.363	0.06

Table 7
TTA at braking onset, planned comparisons to non-assisted driving; FA = false alarms, UA = unnecessary alarms. Larger values indicate a sooner braking onset.

Warning urgency	Warning reliability	Mean difference to baseline [s]	SE	p
Visual-auditory	100%	1.06	0.14	<0.001
	60% with FA	0.93	0.13	<0.001
	60% with UA	1.21	0.13	<0.001
Visual	100%	1.00	0.13	<0.001
	60% with FA	0.93	0.13	<0.001
	60% with UA	0.94	0.13	<0.001

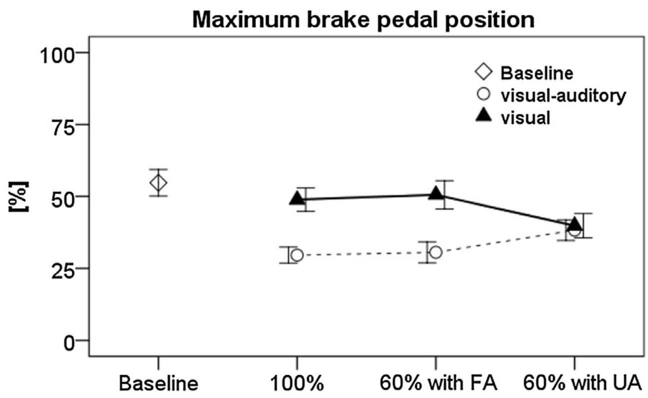


Fig. 5. Maximum brake pedal position, mean and 95%-confidence limits are depicted; FA = false alarms, UA = unnecessary alarms.

Table 8
Maximum brake pedal position, ANOVA results.

Effect	F	df_1	df_2	p	η^2
Warning urgency	9.58	1	53	0.003	0.15
Warning reliability	0.09	2	53	0.912	0.00
Warning reliability * Warning urgency	1.91	2	53	0.159	0.07
Situation	15.15	3	51	<0.001	0.47
Situation * Warning reliability	2.25	6	104	0.044	0.12
Situation * Warning urgency	0.28	3	51	0.841	0.02
Situation * Warning reliability * Warning urgency	1.44	6	104	0.205	0.08

Table 9
Maximum brake pedal position, separate ANOVAs with the between-subjects-factor of warning reliability.

Effect	Situation	F	df_1	df_2	p	η^2
Warning reliability	Turning vehicle	0.32	2	56	0.726	0.01
	Crossing vehicle	0.03	2	57	0.974	0.00
	Crossing pedestrian	0.60	2	57	0.555	0.02
	Crossing cyclist	0.56	2	57	0.577	0.02

maximum brake force, but there was a significant interaction between *warning reliability* and driving *situation* (see Table 8). Separate ANOVAs for each driving situation (see Table 9) with the independent variable *warning reliability* did not reveal any significant effect.

In Table 10, the results of planned comparisons with the baseline drives are shown. The participants applied the brake less forcefully when assisted by the visual-auditory AWS than during non-assisted driving, regardless of the *warning reliability*. In the condition with visual AWS, such a reduction of the maximum brake force was found only in the condition with a 60% reliable AWS with unnecessary alarms (see Fig. 5 and Table 10).

Table 10
Maximum brake pedal position, planned comparisons to non-assisted driving; FA = false alarms, UA = unnecessary alarms. Smaller values indicate a reduction of the maximum brake pedal position.

Warning urgency	Warning reliability	Mean difference to baseline [%]	SE	p
Visual-auditory	100%	−26.35	6.33	<0.001
	60% with FA	−23.59	6.10	<0.001
	60% with UA	−16.44	6.10	0.010
Visual	100%	−5.86	7.54	0.441
	60% with FA	−4.19	7.54	0.581
	60% with UA	−14.89	7.54	0.054

Table 11
Distribution of safety-critical events (SCEs).

Warning urgency	Warning reliability	$N_{validtrials}$	N_{SCE}	$\%_{critical}$
Visual-auditory	100%	120	14	12
	60% with FA	119	23	19
	60% with UA	120	1	1
Visual	100%	120	7	6
	60% with FA	120	11	9
	60% with UA	120	19	16
Baseline	–	240	186	78

Finally, we also assessed whether the heterogeneity of the sample affected the drivers' braking onset and maximum brake pedal position. The ANOVAs were conducted with and without outliers identified in section 2.4. The results of the inferential statistics (i.e., significance level of the factors and interactions) revealed similar data pattern when these outliers were excluded from the analysis.

3.3. Effectiveness

The impact of *warning urgency* and *warning reliability* on the AWS's effectiveness were evaluated using a logistic regression approach. Specifically, a full-factorial logistic regression with the predictors *warning reliability*, *warning urgency* and *driving situation* (same factor levels as in the ANOVA) on the frequency of SCEs was conducted (dependent variable: TTA_{min} less or greater than one second). Each driver contributed three TTA_{min}-values per situation to the analysis. One trial in which the test situation could not be produced was the only one missing from this analysis (see Section 3.1).

First, the frequency of SCEs in the assisted drives were compared (see Table 11 and Table 12). The effect of *warning urgency* was not significant. *Warning reliability* significantly influenced the occurrence of SCEs (main effect *warning reliability*); however, this influence also depended on warning urgency (interaction *warning reliability* * *warning urgency*). Separate logistic regressions per urgency condition showed that false alarms increased the occurrence of SCEs compared with that of the perfectly reliable condition when the warnings were presented visual-auditory (model fit: Wald $X^2 = 51.43$, $df = 11$, $p < 0.001$, contrast: 100% vs. 60% with false alarms: $p = 0.016$, $OR = 4.38$). In contrast, unnecessary warnings even tended to reduce the frequency of SCEs compared with that of the perfectly reliable condition (contrast: 100% vs. 60% with unnecessary alarms: $p = 0.119$, $OR = 0.17$). For visually presented warnings, *warning reliability* did not significantly modulate the number of SCEs (Wald $X^2 = 16.41$, $df = 11$, $p = 0.127$). The driving *situation* did not influence the occurrence of SCEs and there was also no statistically significant three-way interaction with reliability and urgency of the AWS.

Table 12
Logistic regression; dependent variable: $TTA_{min} < 1$ s vs. ≥ 1 s; method: enter; reference categories: visual-auditory (factor: warning urgency), 100% reliable (factor: warning reliability) model fit: $\chi^2 = 67.86$, df = 23, p <0.001.

Predictor	B	Wald	df	p
Warning urgency	−1.76	2.43	1	0.119
Warning reliability	–	12.36	2	0.002
60% with FA vs. 100%	1.48	5.83	1	0.016
60% with UA vs. 100%	−1.76	2.43	1	0.119
Warning reliability*Warning urgency	–	8.89	2	0.012
Situation	–	1.55	3	0.671
Situation * Warning reliability	–	3.75	6	0.710
Situation * Warning urgency	–	2.34	3	0.525
Warning urgency * Situation * Warning reliability	–	0.25	6	1.000

Table 13
Separate logistic regressions for each urgency level; dependent variable: $TTA_{min} < 1$ s vs. ≥ 1 s; method: enter; reference category: Baseline; model fit: visual-auditory: $\chi^2 = 321.13$, df = 3, p <0.001, visual: $\chi^2 = 304.13$, df = 3, p <0.001.

Warning urgency	Predictor	B	Wald	df	p	OR
Visual-auditory	Warning reliability	–	174.95	3	<0.001	–
	100% vs. Baseline	−3.26	101.52	1	<0.001	0.038
	60% with FA vs. Baseline	−2.67	91.34	1	<0.001	0.070
	60% with UA vs. Baseline	−6.02	35.06	1	<0.001	0.002
Visual	Warning reliability	–	208.67	3	<0.001	–
	100% vs. Baseline	−4.02	91.95	1	<0.001	0.018
	60% with FA vs. Baseline	−3.53	100.52	1	<0.001	0.029
	60% with UA vs. Baseline	−2.91	97.81	1	<0.001	0.055

In addition to comparing the assisted drives against each other, they were compared with the baseline drives by two separate logistic regressions for each urgency condition. The results are shown in Table 13. The analysis revealed that the AWS strongly decreased the occurrence of SCEs during the drive, regardless of *warning urgency* and *warning reliability* (see also Table 11).

3.4. Subjective evaluation

The drivers' subjective evaluation of the AWS was assessed during the follow-up interview by using various rating scales. The impact of the warning system's reliability and urgency of the warning output on subjective assessments were analysed via full-factorial 3 × 2 between-subjects ANOVAs with the dependent variables *warning reliability* and *warning urgency* (same factor levels as previously). One missing rating could not be included in the analysis (see Section 3.1). All rating items were subject to a separate ANOVA.

Warning urgency had no significant effect on any of these ratings and there was also no interaction between *warning urgency* and *warning reliability* (see Tables A1 and A2 in the Appendix A). *Warning reliability* influenced the drivers' assessment of the AWS in several rating items, as can be observed in Fig. 6 and Table 14. To prevent inflation of the type I error, the alpha level was adjusted according to Bonferroni (i.e., $p < 0.004$ was required). Planned comparisons to the condition with perfectly reliable AWS are provided

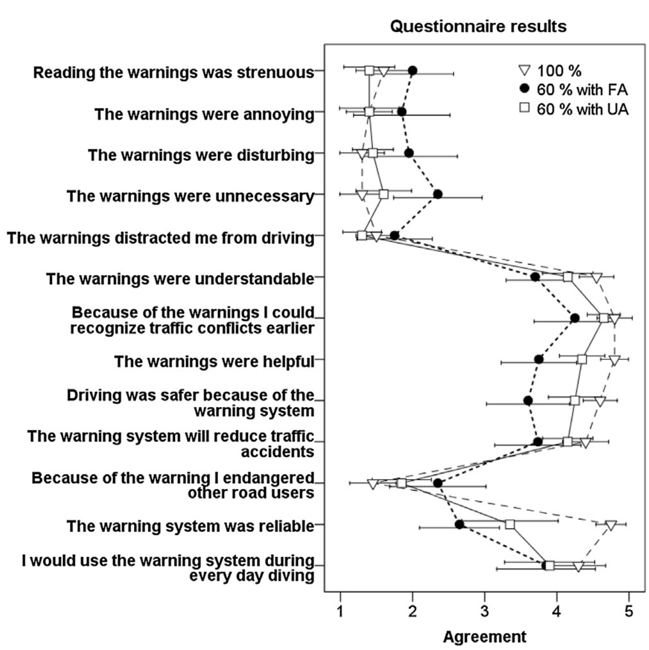


Fig. 6. Questionnaire results collapsed across the urgency conditions, mean and 95%-confidence limits are depicted; FA = false alarms, UA = unnecessary alarms; a five-point scale ranging from "not at all" to "very much" was used.

in Table 15. In the following, we discuss rating items in which a significant effect of *warning reliability* became apparent in detail. However, because of the high correlation between the rating items (see Table A3 in the Appendix A) and because of the high number of simultaneous tests, the results should be treated with caution.

Participants rated the AWS to be less reliable compared with the perfectly reliable AWS both in the condition with false and unnecessary alarms (item: "The warning system was reliable"). In the condition with false alarms, but not in the condition with unnecessary alarms, *ease of use* (items: "The warnings were unnecessary" and "The warnings were understandablë), *usefulness* (items: T̈he warnings were helpful" and "Driving was safer with the warning system") and *trust* (item: "Because of the warning I endangered other road users") decreased in some, but not all of the rating items (see Table 14 and Table 15). The participants' *intention to use* (item "I would use the warning system during every day driving") was not affected by *warning reliability*.

4. Discussion

Using a driving simulator, it was investigated whether false alarms (i.e., without any plausible reason) impact drivers' compliance with a cooperative AWS differently than unnecessary alarms (i.e., with plausible reason because of other traffic participants). During the drive, the participants completed various urban driving situations during which they had to react to traffic conflicts with other road users, such as a cyclist taking the drivers' right of way or a pedestrian entering the road between parked cars. We compared drivers' compliance and effectiveness of a perfectly reliable AWS with that of two 60% reliable warning systems with either false or unnecessary additional alarms and a baseline condition without any assistance. Furthermore, warning urgency was taken into account by comparing a visual-auditory AWS with a purely visual one. The participants reacted faster to the visual-auditory AWS than to the visual one, but they also braked more strongly in the visual warning condition. The impact of warning reliability was dependent on warning urgency. In the visual-auditory condition, compliance with the AWS and warning effectiveness were decreased by false alarms but increased by unnecessary alarms. In contrast, there was no such effect of warning reliability in the condition with visual AWS on both compliance and effectiveness. Compared with that of the baseline drives, an earlier braking onset and a decreased frequency of SCEs was observed even when the AWS issued false alarms. Independent of warning urgency, false alarms decreased drivers' subjective assessments of the warning system (i.e., *usefulness*, *ease of use* and *trust*).

The finding that visual-auditory warnings prompted a sooner braking onset compared with that of purely visual warnings is in line with prior research on driver warnings (e.g., Ho et al., 2007; Kramer et al., 2007) and can be explained by *redundancy gain* (Miller, 1982; Raab, 1962) caused by the additional presentation of the auditory warning component. The warnings used in this study consisted of a visual output that informed the driver of a

Table 14
Follow-up questionnaire, ANOVA results, only effects of "warning reliability" are shown, the complete ANOVA results can be found in the appendix. * = p < 0.004 (significant result after Bonferroni adjustment).

Dimension	Item	F	df_1	df_2	p	η^2
Ease of Use	Reading the warnings was strenuous	2.07	2	54	0.137	0.07
	The warnings were annoying	1.25	2	54	0.295	0.04
	The warnings were disturbing	2.35	2	54	0.105	0.08
	The warnings were unnecessary	**6.49**	**2**	**54**	**0.003**	**0.19***
	The warnings distracted me from driving	1.56	2	54	0.219	0.06
	The warnings were understandable	**6.93**	**2**	**54**	**0.002**	**0.20***
Usefulness	Because of the warnings I could recognize traffic conflicts earlier	2.63	2	54	0.081	0.09
	The warnings were helpful	**9.08**	**2**	**54**	**<.001**	**0.25***
	Driving was safer because of the warning system	**6.28**	**2**	**54**	**0.004**	**0.19***
	The warning system will reduce traffic accidents	2.41	2	53	0.099	0.08
Trust	Because of the warning I endangered other road users	3.64	2	54	0.033	0.12
	The warning system was reliable	**18.40**	**2**	**54**	**<.001**	**0.41***
Intention to use	I would use the warning system during every day driving	0.29	2	54	0.749	0.01

conflict location. This spatial warning was accompanied by an auditory stimulus that did not convey any additional information, a so-called *accessory stimulus* (cf. Kiesel and Miller, 2007). It has been proposed that these accessory stimuli increase the perceived intensity of concurrent stimuli (Stein et al., 1996), speed up cognitive processes involved in response selection (Posner et al., 1976) and influence motor output (Miller et al., 1999). These findings from cognitive psychology are in line with the result from the current study that drivers responded more rapidly in the visual-auditory condition. Thus, the goal of designing an AWS with a more-urgent and a less-urgent warning was accomplished, as warning urgency is associated with faster reactions (Politis et al., 2013). There was no significant main effect of warning urgency on the frequency of SCEs, which can be explained by the stronger brake reaction in the visual warning condition. Thus, it appears that the participants compensated the later braking onset by braking harder, so that warning effectiveness was ultimately not decreased when the warnings were only presented visually. Such compensatory braking behaviour has also been reported by Naujoks et al. (2015b) in the context of partially automated driving.

In the condition with visual-auditory AWS, false alarms caused a delayed response to the AWS that resulted in a higher frequency of SCEs compared with that of the perfectly reliable AWS condition. Taken together, these results noted the occurrence of the *cry wolf effect* (Breznitz, 1984) in the visual-auditory condition. It has been suggested that the reason for a decreased compliance may be the diminished contingency between the presentation of driver warnings and the need for action (Kiesel and Miller, 2007). Unnecessary warnings sped up the drivers' response to the AWS and, in tendency, the effectiveness of the warning system was increased compared with that of the perfectly reliable AWS. These results are in line with the assumption that the drivers' compliance is only impaired by "truly" false alarms that are presented without an obvious reason (Cotté et al., 2001; Lees and Lee, 2007). The differential impact of false and unnecessary alarms on drivers' compliance has theoretically been explained by assuming that these warnings put the driver in a position to better understand the *process* involved in the generation of a warning (Lees and Lee, 2007; Lee and See, 2004). Following this explanation, drivers in the condition with unnecessary warnings might have concluded that the AWS detected other road users that were likely, but not certain to cross the host vehicle's path. By observing the test situations, they might thus have formed a mental model of how warnings are generated by the AWS. Drivers in the conditions with false alarms were not provided with this possibility as false alarms were always presented without any plausible reason. With regard to the generalisability of these findings, it is thus important to note that the differential effects of false and unnecessary alarms in this study are at least partly attributable to the situational context that made it either possible or impossible to understand the *process* involved in the generation of warnings. Other possibilities to manipulate this trust dimension, like informing drivers about the *process* of the AWS (e.g., by a technical description of the AWS), were not part of the study.

The most important finding from this study, which has not been reported before, is that warning reliability did not influence compliance with the AWS when warnings were presented only visually. There was also no significant effect of warning reliability on the effectiveness of the warning system in the visual AWS condition. These findings raise the question of why false and unnecessary alarms did not cause the same consequences on compliance as in the visual-auditory condition. Following the explanatory framework presented by Lees and Lee (2007), a *cry wolf effect* should have also been present in the visual AWS condition with false alarms. It could be speculated that the participants did not interpret false warnings to be unjustified when they were presented less urgently, which may have prevented the diminished contingency between the presentation of a warning and the need for action. Consequently, the findings support the idea that the less urgent warnings are the more "pardonable" from a driver's perspective (Hoffmann and Gayko, 2012). Furthermore, the result may be due

Table 15
Planned contrasts with the 100% reliability condition. Items are only included in the table if a significant effect of warning reliability was observed. * = p < 0.004 (significant result after Bonferroni adjustment).

Item	Planned contrasts			
	100% vs. 60% with FA		100% vs. 60% with UA	
	mean difference	p	mean difference	p
The warnings were unnecessary	**1.05**	**0.001***	0.30	0.332
The warnings were understandable	**−0.85**	**<0.001***	−0.40	0.085
The warnings were helpful	**−1.05**	**<0.001***	−0.45	0.077
Driving was safer because of the warning system	**−1.00**	**0.001***	−0.35	0.220
The warning system was reliable	**−2.10**	**<0.001***	−1.40	<0.001

to the mode of presentation. In this study, a simulated HUD was used to present warnings, which facilitated the interpretation of the warnings without the need to take one's eyes off the road. It could be possible that the absence of a *cry wolf effect* in the visual AWS condition was partly due to the low costs of paying attention to the warnings when they were false (Wickens and McCarley, 2007). For example, it may have been possible that the *cry wolf effect* would have been observable if the visual warnings required drivers to take their eyes off the road or if they would have been harder to interpret. This theoretical question clearly needs further research from more-controlled research environments so that it can be fully answered.

Regarding the comparisons made to the baseline drives, it is important to note that, in spite of the reduced compliance with and effectiveness of the AWS by false alarms, there was still a large difference compared with the manual drives. Namely, even when the reliability of the AWS was as low as 60%, drivers were still effectively assisted. From a practical perspective, this result reduces the concern that an increased rate of false alarms, due to the predictive nature of cooperative warning systems, will eliminate their potentially positive effects on driving safety. In contrast, the subjective data obtained in the study showed a decrease in the perceived *ease of use*, *usefulness* and *trust* of the AWS, although the reliability of the AWS did not influence the participants' agreement with the statement, "I would use the warning system during every day driving".

From a methodological perspective, one important finding of the study is that the interaction effect between warning reliability and urgency was not reflected in the participants' subjective assessments of the AWS. While behavioural effects of false alarms were evident only in the visual-auditory condition, we found that false alarms negatively affected the participants' evaluation of the AWS's *usefulness*, *ease of use* and *trust*, regardless of warning urgency. One possible explanation for this finding may be that the participants filled out the questionnaires after completion of the whole test drive. Interviewing the drivers directly after each of the driving situations might have led to a greater accordance of behavioural and subjective measures (e.g., Naujoks and Neukum, 2014a). However, from the findings of this study, it seems questionable to rely solely on either subjective or behavioural measures when evaluating driver warning systems. Note that these results should be treated with caution as the survey contained many rating items. Apart from correlations between those items, the high number of statistical tests raises the problem of type I error inflation. In this regard, the survey methodology should be improved in future studies.

Finally, it has to be emphasized that the results obtained in this simulator study cannot be easily translated to real-world driving for several reasons. For example, the frequency of false, unnecessary and correct warnings of real-world cooperative warning systems are not known. It might well be that warnings, either correct or unnecessary/false, would be experienced with a much lower frequency than the warnings investigated in this study (e.g., one per drive or one per week). The ratio of correct and false/unnecessary alarms could also be drastically different than the one investigated, which might prevent the *cry wolf effect* in the first place. Long-term experience with unreliable cooperative warning systems, which cannot be observed in a single-visit simulator drive, might also change the way that drivers respond to false/unnecessary warnings. It is well possible that the impact of false and unnecessary alarms on attitudes and behaviour magnifies over time and might be particularly large in discretionary warning systems that can be turned off. In the first few drives, drivers might choose to turn off a system prone to false alarms, but might continue to use a system with unnecessary alarms. This could dramatically affect the differential benefit of warning systems that are prone to false or unnecessary alarms in the long term. Lastly, it should be noted that although the sample was quite heterogeneous with regard to age, driving experience and mileage, special groups of drivers like novice drivers or elderly were not part of the study. It is possible that these drivers will respond differently to false and unnecessary warnings so that the findings of the current study cannot be easily generalised to them.

5. Conclusions

This study investigated conditions that lead to impaired compliance with and effectiveness of a cooperative advisory warning system (AWS). The following conclusions can be drawn from the study:

1. False alarms (alarms without any obvious reason), but not unnecessary alarms (alarms with another road user present that could have caused the warning), led to decreased compliance with and effectiveness of the AWS depending on warning urgency. Warning reliability only influenced compliance in cases in which the advisory warnings were presented in a more-urgent visual-auditory way but not when presented in a less-urgent visual way.
2. Subjective evaluations of the AWS (e.g., *ease of use*, *usefulness*, *trust*) were negatively influenced by false alarms but not by unnecessary alarms, independent of warning urgency.
3. Compared with a baseline drive without any assistance, the drivers benefited from the AWS, even among cases with false alarms, as they still responded much earlier to critical driving situations.

Acknowledgements

This study was part of the joint Ko-PER research project within the scope of the Ko-FAS Research Initiative (www.kofas.de) under support code 19S9022C. It was sponsored by the German Federal Ministry of Economics and Technology.

Appendix A.

See Tables A1–A3.

Table A1
Descriptive statistics of follow-up interview for each reliability condition, agreement scale ranging from 1 to 5.

Item	Reliability	Mean	SE	N_{valid}
Reading the warnings was strenuous	100%	1.60	0.18	20
	60% with FA	2.00	0.27	20
	60% with UA	1.40	0.17	20
The warnings were annoying	100%	1.40	0.15	20
	60% with FA	1.85	0.32	20
	60% with UA	1.40	0.20	20
The warnings were disturbing	100%	1.30	0.15	20
	60% with FA	1.95	0.32	20

Table A1 (*Continued*)

Item	Reliability	Mean	SE	N_{valid}
	60% with UA	1.45	0.14	20
The warnings were unnecessary	100%	1.30	0.15	20
	60% with FA	2.35	0.29	20
	60% with UA	1.60	0.18	20
The warnings distracted me from driving	100%	1.50	0.14	20
	60% with FA	1.75	0.25	20
	60% with UA	1.30	0.13	20
The warnings were understandable	100%	4.55	0.11	20
	60% with FA	3.70	0.19	20
	60% with UA	4.15	0.17	20
Because of the warnings I could recognize traffic conflicts earlier	100%	4.80	0.12	20
	60% with FA	4.25	0.27	20
	60% with UA	4.65	0.11	20
The warnings were helpful	100%	4.80	0.09	20
	60% with FA	3.75	0.25	20
	60% with UA	4.35	0.15	20
Driving was safer because of the warning system	100%	4.60	0.11	20
	60% with FA	3.60	0.28	20
	60% with UA	4.25	0.18	20
The warning system will reduce traffic accidents	100%	4.40	0.15	20
	60% with FA	3.74	0.28	19
	60% with UA	4.15	0.17	20
Because of the warning I endangered other road users	100%	1.45	0.15	20
	60% with FA	2.35	0.32	20
	60% with UA	1.85	0.20	20
The warning system was reliable	100%	4.75	0.10	20
	60% with FA	2.65	0.26	20
	60% with UA	3.35	0.32	20
I would use the warning system during every day driving	100%	4.30	0.19	20
	60% with FA	3.85	0.33	20
	60% with UA	3.90	0.30	20

Table A2
ANOVA results of the questionnaire.

Dimension	Item	Effect	F	df_1	df_2	p	η^2
Ease of Use	Reading the warnings was strenuous	Warning urgency	0	1	54	1	0
		Warning reliability	2.07	2	54	0.137	0.07
		Interaction	1.55	2	54	0.222	0.05
	The warnings were annoying	Warning urgency	0.14	1	54	0.711	0
		Warning reliability	1.25	2	54	0.295	0.04
		Interaction	1.71	2	54	0.19	0.06
	The warnings were disturbing	Warning urgency	0.00	1	54	1	0
		Warning reliability	2.35	2	54	0.105	0.08
		Interaction	0.46	2	54	0.636	0.02
	The warnings were unnecessary	Warning urgency	0.46	1	54	0.5	0.01
		Warning reliability	6.49	2	54	0.003	0.19
		Interaction	2.46	2	54	0.095	0.08
	The warnings distracted me from driving	Warning urgency	1.26	1	54	0.267	0.02
		Warning reliability	1.56	2	54	0.219	0.06
		Interaction	0.8	2	54	0.457	0.03
	The warnings were understandable	Warning urgency	0.13	1	54	0.722	0
		Warning reliability	6.93	2	54	0.002	0.2
		Interaction	1.37	2	54	0.262	0.05
Usefulness	Because of the warnings I could recognize traffic conflicts earlier	Warning urgency	2.71	1	54	0.105	0.05
		Warning reliability	2.63	2	54	0.081	0.09
		Interaction	2.14	2	54	0.127	0.07
	The warnings were helpful	Warning urgency	0	1	54	1	0
		Warning reliability	9.08	2	54	<0.001	0.25
		Interaction	2.05	2	54	0.139	0.07
	Driving was safer because of the warning system	Warning urgency	0.51	1	54	0.479	0.01
		Warning reliability	6.28	2	54	0.004	0.19
		Interaction	0.39	2	54	0.682	0.01
	The warning system will reduce traffic accidents	Warning urgency	0.3	1	53	0.584	0.01
		Warning reliability	2.41	2	53	0.099	0.08
		Interaction	0.58	2	53	0.565	0.02

Table A2 (*Continued*)

Dimension	Item	Effect	F	df_1	df_2	p	η^2
Trust	Because of the warning I endangered other road users	Warning urgency	0.73	1	54	0.396	0.01
		Warning reliability	3.64	2	54	0.033	0.12
		Interaction	0.42	2	54	0.661	0.02
	The warning system was reliable	Warning urgency	0.12	1	54	0.73	0
		Warning reliability	18.4	2	54	<0.001	0.41
		Interaction	0.64	2	54	0.529	0.02
Intention to use	I would use the warning system during every day driving	Warning urgency	0.76	1	54	0.388	0.02
		Warning reliability	0.29	2	54	0.749	0.01
		Interaction	0.49	2	54	0.614	0.02

Table A3
Correlation matrix of the questionnaire items. Pearson correlations are used. * = p < 0.05. Items are numbered according to Table 4.

Item		2	3	4	5	6	7	8	9	10	11	12	13
1	r	0.484*	0.497*	0.300*	.589*	−0.273*	−0.330*	−0.335*	−0.360*	−0.171	0.267*	−0.193	−0.323*
	p	<0.001	<0.001	0.020	<0.001	0.035	0.010	0.009	0.005	0.196	0.039	0.140	0.012
2	r	1	0.896*	0.726*	0.655*	−0.274*	−0.734*	−0.635*	−0.582*	−0.508*	0.502*	−0.235	−0.679*
	p		<0.001	<0.001	<0.001	0.034	<0.001	<0.001	<0.001	<0.001	<0.001	0.071	<0.001
3	r		1	0.783*	0.615*	−0.376*	−0.782*	−0.692*	−0.649*	−0.538*	0.622*	−0.327*	−0.603*
	p			<0.001	<0.001	0.003	<0.001	<0.001	<0.001	<0.001	<0.001	0.011	<0.001
4	r			1	0.451*	−0.387*	−0.689*	−0.680*	−0.676*	−0.630*	0.506*	−0.431*	−0.548*
	p				<0.001	0.002	<0.001	<0.001	<0.001	<0.001	<0.001	0.001	<0.001
5	r				1	−0.425*	−0.416*	−0.453*	−0.529*	−0.385*	0.337*	−0.271*	−0.331*
	p					0.001	0.001	<0.001	<0.001	0.003	0.009	0.037	0.010
6	r					1	0.373*	0.472*	0.613*	0.441*	−0.433*	0.666*	0.312*
	p						0.003	<0.001	<0.001	<0.001	0.001	<0.001	0.015
7	r						1	0.822*	0.691*	0.562*	−0.543*	0.310*	0.655*
	p							<0.001	<0.001	<0.001	<0.001	0.016	<0.001
8	r							1	.752*	0.535*	−0.575*	0.513*	0.570*
	p								<0.001	<0.001	<0.001	<0.001	<0.001
9	r								1	0.813*	−0.607*	0.636*	0.623*
	p									<0.001	<0.001	<0.001	<0.001
10	r									1	−0.358*	0.541*	0.590*
	p										0.005	<0.001	<0.001
11	r										1	−0.378*	−0.416*
	p											0.003	0.001
12	r											1	0.301*
	p												0.019

References

Aust, M.L., Engström, J., Viström, M., 2013. Effects of forward collision warning and repeated event exposure on emergency braking. Transp. Res. Part F 18, 34–46.

Bliss, J.P., Acton, S.A., 2003. Alarm mistrust in automobiles: how collision alarm reliability affects driving. Appl. Ergon. 34 (6), 499–509.

Breznitz, S., 1984. Cry-wolf: The Psychology of False Alarms. Lawrence Erlbaum Associates, Hillsdale, NJ.

Campbell, J.L., Richard, C.M., Brown, J.L., McCallum, M., 2007. Crash Warning System Interfaces: Human Factors Insights and Lessons Learned (No. HS 810 697). National Highway Traffic Safety Administration, Washington, D.C.

Cotté, N., Meyer, J., Coughlin, J.F., 2001. Older and younger drivers' reliance on collision warning systems. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 45 (4), 277–280.

Cummings, M.L., Kilgore, R.M., Wang, E., Tijerina, L., Kochhar, D.S., 2007. Effects of single versus multiple warnings on driver performance. Hum. Factors 49 (6), 1097–1106.

Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. 13 (3), 319–340.

Dzindolet, M.T., Peterson, S.A., Pomranky, R.A., Pierce, L.G., Beck, H.P., 2003. The role of trust in automation reliance. Int. J. Hum. Comput. Stud. 58 (6), 697–718.

Edquist, J., Rudin-Brown, C.M., Lenné, M.G., 2012. The effects of on-street parking and road environment visual complexity on travel speed and reaction time. Acc. Anal. Prev. 45, 759–765.

Federal Statistical Office, 2013. Verkehrsunfälle [Traffic Accidents] (Fachserie 8, Reihe 7). Federal Statistical Office, Wiesbaden.

Finger, R., Bisantz, A.M., 2002. Utilizing graphical formats to convey uncertainty in a decision-making task. Theoretical Issues in Ergonomics Science 3 (1), 1–25.

Getty, D.J., Swets, J.A., Pickett, R.M., Gonthier, D., 1995. System operator response to warnings of danger: a laboratory investigation of the effects of the predictive value of a warning on human response time. J. Exp. Psychol.: Appl. 1 (1), 19–33.

Gold, C., Körber, M., Hohenberger, C., Lechner, D., Bengler, K., 2015. Trust in automation–before and after the experience of take-over scenarios in a highly automated vehicle. Procedia Manuf. 3, 3025–3032.

Gupta, N., Bisantz, A.M., Singh, T., 2002. The effects of adverse condition warning system characteristics on driver performance: an investigation of alarm signal type and threshold level. Behav. Inf. Technol. 21 (4), 235–248.

Hamdar, S.H., Qin, L., Talebpour, A., 2016. Weather and road geometry impact on longitudinal driving behavior: exploratory analysis using an empirically supported acceleration modeling framework. Transp. Res. Part C 67, 193–213.

Hayward, J.C., 1972. Near miss determination through use of a scale of danger. Highway Res. Rec. 384, 24–34.

Ho, C., Reed, N., Spence, C., 2007. Multisensory in-car warning signals for collision avoidance. Hum. Factors 49 (6), 1107–1114.

Hoffmann, S., Krüger, H.-P., Buld, S., 2003. Vermeidung von simulator sickness anhand eines trainings zur gewöhnung an die fahrsimulation. In: VDI-Gesellschaft Fahrzeug- und Verkehrstechnik (Ed.), Simulation Und Simulatoren – Mobilität Virtuell Gestalten (VDI-Berichte, Nr. 1745, pp. 385–404). VDI Verlag, Düsseldorf.

Hoffmann, Gayko, J.E., 2012. Fahrerwarnelemente. In: Winner, H., Hakuli, S., Wolf, G. (Eds.), Handbuch Fahrerassistenzsysteme. Vieweg & Teubner, Wiesbaden, pp. 343–354.

ISO 15623:2013 (E), 2013. Intelligent Transport Systems – Forward Vehicle Collision Warning Systems – Performance Requirements and Test Procedures. ISO, Genf.

Jian, J.-Y., Bisantz, A.M., Drury, C.G., 2000. Foundations for an empirically determined scale of trust in automated systems. Int. J. Cognit. Ergon. 4 (1), 53–71.

Kestin, I.G., Miller, B.R., Lockhart, C.H., 1988. Auditory alarms during anesthesia monitoring. Anesthesiology 69 (1), 106–108.

Kiesel, A., Miller, J., 2007. Impact of contingency manipulations on accessory stimulus effects. Percept. Psychophys. 69 (7), 1117–1125.

Kragt, H., Bonten, J., 1983. Evaluation of a conventional process-alarm system in a fertilizer plant. IEEE Trans. Syst. Man Cybern. 13 (4), 586–600.

Kramer, A.F., Cassavaugh, N., Horrey, W.J., Becic, E., Mayhugh, J.L., 2007. Influence of age and proximity warning devices on collision avoidance in simulated driving. Hum. Factors 49 (5), 935–949.

Lee, J.D., Hoffman, J.D., Hayes, E., 2004. Collision warning design to mitigate driver distraction. In: Herbsleb, J., Olson, G. (Eds.), Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, New York, pp. 65–72.

Lee, J.D., Moray, N., 1992. Trust, control strategies and allocation of function in human- machine systems. Ergonomics 35 (10), 1243–1270.

Lee, J.D., See, K.A., 2004. Trust in automation: designing for appropriate reliance. Hum. Factors 46 (1), 50–80.

Lees, M.N., Lee, J.D., 2007. The influence of distraction and driving context on driver response to imperfect collision warning systems. Ergonomics 50 (8), 1264–1286.

Lenné, M.G., Triggs, T.J., Mulvihill, C.M., Regan, M.A., Corben, B.F., 2008. Detection of emergency vehicles: driver responses to advance warning in a driving simulator. Hum. Factors 50 (1), 135–144.

Lenné, M.G., Triggs, T.J., 2009. Warning drivers of approaching hazards: the importance of location clues and multi-sensory cues. In: de Waard, D., Godthelp, J., Kooi, F.L., Brookhuis, K.A. (Eds.), Human Factors, Security and Safety. Shaker Publishing, Maastricht, pp. 203–211.

Maag, C., Schneider, N., Lübbeke, T., Weisswange, T.H., Goerick, C., 2015. Car Gestures–Advisory warning using additional steering wheel angles. Acc. Anal. Prev. 83, 143–153.

Marciano, H., Yeshurun, Y., 2015. Perceptual load in different regions of the visual scene and its relevance for driving. Hum. Factors 57 (4), 701–716.

Mahr, A., Cao, Y., Theune, M., Dimitrova-Krause, V., Schwartz, T., Müller, C., 2010. What if it suddenly fails – Behavioral aspects of advanced driver assistant systems on the example of local danger alerts. In: Coelho, H., Studer, R., Wooldridge, M. (Eds.), Proceedings of 19th European Conference on Artificial Intelligence (ECAI 2010) – Frontiers in Artificial Intelligence and Applications. IOS Press, Lissabon, pp. 1051–1052.

Meredith, C., Edworthy, J., 1995. Are there too many alarms in the intensive care unit? An overview of the problems. J. Adv. Nurs. 21 (1), 15–20.

Meyer, J., 2004. Conceptual issues in the study of dynamic hazard warnings. Hum. Factors 46 (2), 196–204.

Miller, J.O., 1982. Divided attention: evidence for coactivation with redundant signals. Cognit. Psychol. 14 (2), 247–279.

Miller, J., Franz, V., Ulrich, R., 1999. Effects of auditory stimulus intensity on response force in simple, go/no-go, and choice RT tasks. Percept. Psychophy. 61 (1), 107–119.

Montgomery, D.A., 1999. Human sensitivity to variability information in detection decisions. Hum. Factors 41 (1), 90–105.

Montgomery, D.A., Sorkin, R.D., 1996. Observer sensitivity to element reliability in a multielement visual display. Hum. Factors 38 (3), 484–494.

Naujoks, F., Grattenthaler, H., Neukum, A., Weidl, G., Petrich, D., 2015a. Effectiveness of advisory warnings based on cooperative perception. IET Intel. Transport Syst. 9 (6), 606–617.

Naujoks, F., Neukum, A., 2014a. Timing of in-vehicle advisory warnings based on cooperative perception. In: de Waard, K., Brookhuis, R., di Nocera, R., Brouwer, P., Barham, C., Weikert, A., Kluge, W., Gerbino, A. (Eds.), Proceedings of the Human Factors and Ergonomics Society Europe Chapter Annual Conference (pp. 193–206, Online).

Naujoks, F., Neukum, A., 2014b. Specificity and timing of advisory warnings based on cooperative perception. In: Butz, A., Koch, M., Schlichter, J. (Eds.), Mensch & Computer 2014–Workshopband (S. 229–238). De Gruyter Oldenbourg, Berlin.

Naujoks, F., Purucker, C., Neukum, A., Wolter, S., Steiger, R., 2015b. Controllability of partially automated driving functions – does it matter whether drivers are allowed to take their hands off the steering wheel? Transp. Res. Part F 35, 185–198.

Naujoks, F., Totzke, I., 2014. Behavioral adaptation caused by predictive warning systems – the case of congestion tail warnings. Transp. Res. Part F 26, 49–61.

Politis, I., Brewster, S., Pollick, F., 2013. Evaluating multimodal driver displays of varying urgency. In: Paper Presented at the Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. Eindhoven, the Netherlands.

Posner, M.I., Nissen, M.J., Klein, R.M., 1976. Visual dominance: an information-processing account of its origins and significance. Psychol. Rev. 83 (2), 157–171.

Pritchett, A.R., 2001. Reviewing the role of cockpit alerting systems. Hum. Factors Aerosp. Saf. 1 (1), 5–38.

Raab, D.H., 1962. Statistical facilitation of simple reaction times. Trans. N. Y. Acad. Sci. 24 (5), 574–590.

Rogé, J., Douissembekov, E., Vienne, F., 2012. Low conspicuity of motorcycles for car drivers: dominant role of bottom-up control of visual attention or deficit of top-down control? Hum. Factors 54 (1), 14–25.

Roulston, M.S., Smith, L.A., 2004. The boy who cried wolf revisited: the impact of false alarm intolerance on cost-loss scenarios. Weather Forecasting 19 (2), 391–397.

Seeliger, F., Weidl, G., Petrich, D., Naujoks, F., Breuel, G., Neukum, A., Dietmayer, K., 2014. Advisory warnings based on cooperative perception. In: Institute of Electrical, Electronics Engineers, I.E.E.E. (Eds.), Proceedings of the IEEE Intelligent Vehicles Symposium (pp. 246–252). IEEE, Dearborn.

Shah, S.J., Bliss, J.P., Chancey, E.T., Brill, J.C., 2015. Effects of alarm modality and alarm reliability on workload, trust, and driving performance. Proc. Hum. Factors Ergon. Soc. Ann. Meet. 59 (1), 1535–1539.

Sorkin, R.D., 1988. Why are people turning off our alarms? J. Acoust. Soc. Am. 84 (3), 1107–1108.

Sorkin, R.D., Kantowitz, B.H., Kantowitz, S.C., 1988. Likelihood alarm displays. Hum. Factors 30 (4), 445–459.

Stein, B.E., London, N., Wilkinson, L.K., Price, D.D., 1996. Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. J. Cogn. Neurosci. 8 (6), 497–506.

Totzke, I., Naujoks, F., Mühlbacher, D., Krüger, H.-P., 2012. Precision of congestion warnings: do drivers really need warnings with precise information about the congestion tail's position? In: de Waard, D., Merat, N., Jamson, A.H., Barnard, Y., Carsten, O.M.Y. (Eds.), Human Factors of Systems and Technology (pp 235–248). Shaker Publishing, Maastricht.

Tukey, J.W., 1977. Exploratory Data Analysis. Addison-Wesley Publishing Company, Reading, UK.

United States Department of Transportation, 2015. U.S. Department of Transportation Announces up to $42 Million in Next Generation Connected Vehicle Technologies. United States Department of Transportation (Retrieved on September first, 2016. URL: http://www.its.dot.gov/press/2015/ngv_tech_announcement.htm).

van der Horst, A.R.A., 1990. A Time-based Analysis of Road User Behaviour in Normal and Critical Encounters. Technische Universiteit Delft, Dissertation.

Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D., 2003. User acceptance of information technology: toward a unified view. MIS Q. 27 (3), 425–478.

Weidl, G., Breuel, G., 2012. Overall probabilistic framework for modeling and analysis of intersection situations. In: Paper Presented at the 16th Int. Forum on Advanced Microsystems for Automotive Applications (AMAA 2012), Berlin, Germany.

Weidl, G., Singhal, V., Petrich, D., Kasper, D., Wedel, A., Breuel, G., 2013. Collision risk prediction and warning at road intersections using an object oriented Bayesian network. In: Paper Presented at the 5th International Conference Automotive-UI13, Eindhoven, The Netherlands.

Weller, G., Heyne, F., Feige, T., Bretschneider, H., Oeser, H., Schlag, B., 2013. Die wirkung gerichteter warnungen von fahrerassistenzsystemen auf die blickzuwendungs- und reaktionszeiten von autofahrern. In: Brandenburg, E., Doria, L., Gross, A., Günzler, T., Smieszek, H. (Eds.), Grundlagen Und Anwendungen Der Mensch-Maschine-Interaktion: 10. Berliner Werkstatt Mensch-Maschine-Systeme. Universitätsverlag der TU Berlin, Berlin, pp. 376–382.

Werneke, J., Kleen, A., Vollrath, M., 2013. Perfect timing – urgency, not driving situations, influence the best timing to activate warnings. Hum. Factors 56 (2), 249–259.

Wickens, C.D., McCarley, J.S., 2007. Applied Attention Theory. Taylor & Francis.

Wickens, C.D., Dixon, S.R., 2007. The benefits of imperfect diagnostic automation: a synthesis of the literature. Theor. Issues Ergon. Sci. 8 (3), 201–212.

Winkler, S., Werneke, J., Vollrath, M., 2016. Timing of early warning stages in a multi stage collision warning system: drivers' evaluation depending on situational influences. Transp. Res. Part F 36, 57–68.

Woods, D.D., 1995. The alarm problem and directed attention in dynamic fault management. Ergonomics 38 (11), 2371–2393.

Yamada, K., Kuchar, J.K., 2006. Preliminary study of behavioral and safety effects of driver dependence on a warning system in a driving simulator. IEEE Trans. Syst. Man Cybern. Part A 36 (3), 602–610.

Zarife, R., 2014. Integrative Warning Concept for Multiple Driver Assistance Systems. Dissertation. Universität Würzburg.