Accident Analysis and Prevention 101 (2017) 55–66

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Full length article

Augmented reality warnings in vehicles: Effects of modality and specificity on effectiveness

Felix Schwarz[a,*], Wolfgang Fastenmeier[b]

[a] Psychologische Hochschule Berlin, Theresienstr. 48, D-80333 München, Germany
[b] Psychologische Hochschule Berlin, Hochkönigstr. 6, D-81825 München, Germany

ARTICLE INFO

Article history:
Received 24 June 2016
Received in revised form 23 January 2017
Accepted 26 January 2017
Available online 9 February 2017

Keywords:
Collision avoidance
Sight obstruction
Reliability
Warning design
Specific warnings
Augmented reality
Spatial referencing
Specificity

ABSTRACT

In the future, vehicles will be able to warn drivers of hidden dangers before they are visible. Specific warning information about these hazards could improve drivers' reactions and the warning effectiveness, but could also impair them, for example, by additional cognitive-processing costs.

In a driving simulator study with 88 participants, we investigated the effects of modality (auditory vs. visual) and specificity (low vs. high) on warning effectiveness. For the specific warnings, we used augmented reality as an advanced technology to display the additional auditory or visual warning information. Part one of the study concentrates on the effectiveness of necessary warnings and part two on the drivers' compliance despite false alarms.

For the first warning scenario, we found several positive main effects of specificity. However, subsequent effects of specificity were moderated by the modality of the warnings. The specific visual warnings were observed to have advantages over the three other warning designs concerning gaze and braking reaction times, passing speeds and collision rates. Besides the true alarms, braking reaction times as well as subjective evaluation after these warnings were still improved despite false alarms. The specific auditory warnings were revealed to have only a few advantages, but also several disadvantages.

The results further indicate that the exact coding of additional information, beyond its mere amount and modality, plays an important role. Moreover, the observed advantages of the specific visual warnings highlight the potential benefit of augmented reality coding to improve future collision warnings.

1. Introduction

Sight obstruction is one of the most relevant contributing factors for causing accidents in road traffic (Staubach, 2009). Current collision-avoidance systems reach their limits in corresponding scenarios. New technologies like connecting road users via wireless networks (Car-to-X) will enable future systems to detect dangers earlier and even when they are occluded by obstacles like other vehicles or buildings (Fuerstenberg et al., 2007; Pierowicz et al., 2000; Seeliger et al., 2014). Therefore, warnings of such dangers should provide a huge potential for improving safety (Naujoks et al., 2014).

However, the introduction of additional warnings in vehicles may unintentionally confront drivers with a rising frequency of false alarms. The generally limited reliability of automatic accident prediction (Lees, 2010; Parasuraman et al., 1997; Zabyshny and Ragland, 2003) is likely to further decrease due to earlier output, the higher complexity of the assisted situations (Hägglund, 2008) as well as limitations of the technology like latencies of data transmission (Lu et al., 2005).

False alarms are prone to annoy drivers by unnecessarily capturing their attention. This can lower drivers' acceptance of the system (LeBlanc and Tsimhoni, 2008) in terms of a low willingness to use it (Dillon, 2001). The "cry-wolf syndrome" (Breznitz, 1984) describes the phenomenon that frequent false alarms lower operators' trust in a system (Parasuraman and Riley, 1997). Behavioral consequences of annoyance (Kiefer et al., 1999; Lerner et al., 1996), as well as distrust, include slower and weaker braking responses or even a tendency to ignore or turn off warnings after several false alarms (Bliss et al., 1995; Chugh and Caird, 1999; Getty et al., 1995; Sorkin et al., 1988). While annoyance is defined as a subjective response, that has been used mostly in relation to acoustic stimuli (e.g. Marshall et al., 2007), trust is defined as an attitude about the utility of an agent to reach set goals (Lee and See, 2004). Overall, the potential negative consequences of frequent false alarms like

* Corresponding author.
E-mail addresses: schwarzfelix@gmail.com (F. Schwarz), w.fastenmeier@psychologische-hochschule.de (W. Fastenmeier).

http://dx.doi.org/10.1016/j.aap.2017.01.019
0001-4575/

annoyance, loss of acceptance, and loss of trust should get increased attention when evaluating long-term effectiveness of future warning systems.

The combination of false alarms and occluded hazards could be even more problematic in that respect. Depending on the comprehension of the cause of a false alarm, Lees and Lee (2007) proposed the distinction between unnecessary (user can understand the cause) and false warnings (user cannot understand the cause). While unnecessary warnings can even support trust (Maltz and Shinar, 2004, 2007), false warnings have mostly negative consequences (Lees and Lee, 2007). Sight obstructions can impair the detection of the cause of a warning so that an unnecessary warning is perceived as a false warning. For example, a driver who receives a warning and sees a pedestrian running towards the road but stopping just in time will probably understand the cause of this warning and perhaps even learn to better assess the system. In the case where that same pedestrian has been occluded by parked vehicles, the driver probably would not understand the warning cause, might assess it as a false warning, lose trust in the system, and ignore the next warning. Thus, sight obstruction of hazards could amplify the negative effects of false alarms.

1.1. Warning design

There is an extensive body of research and guidelines for the design of the human-machine interface (HMI) of collision warnings (Campbell et al., 2007; COMSIS Corporation, 1996; Green et al., 1995; Informal Group on Intelligent and Transport Systems, 2011). However, nearly all of them refer to warnings that are issued in situations where the hazard is directly visible to the driver. For warnings of hazards that are hidden at the time of the warning onset, there are changes in some fundamental circumstances like available information or rates of false alarms, and it is very unclear how to optimally design such warnings (Naujoks et al., 2014). The goal of these warnings is still to support drivers in avoiding potential collisions. Because of cognitive and behavioral preconditions for the drivers to succeed, however, it might not be optimal to just elicit or guide their attention. Conveying specific warning information to support drivers' situation awareness, despite sight obstructions, could be decisive to enable drivers to rapidly select and execute the optimal response to an actual hazard. At the same time, it will be crucial to ensure low annoyance and appropriate trust in order to preserve compliance despite false alarms. Accordingly, some general insights about warning designs should be reconsidered in the light of the possibilities and limitations of Car-to-X warnings.

1.2. Specificity of warnings

Warnings can contain different amounts of specific information about a hazard, for example, its position and motion direction or type. Immediately after a warning of a hidden danger, the only available information for a driver is the traffic environment and the warning message itself. Obviously, specific warning information needs to become more relevant to improve drivers' reactions in corresponding situations.

Based on the model of stages of warning information processing (cf. Wogalter, 2006), more specific warnings can theoretically improve the comprehension of the cause of the warning and cause or accelerate the attention allocation towards the location of the hazard. Referring to the construct of situation awareness (Endsley, 1995), additional specific warning information can support all three of its components: (1) perception of the specified feature of the otherwise hidden opponent, (2) comprehension of the cause of the warning, as well as (3) projection of the appearance of an opponent. This in turn could materialize in faster hazard detection (as soon as it becomes visible), quicker and stronger braking reactions, and a generally more efficient collision mitigation after true alarms. In addition, there could be reduced deterioration of compliance after false alarms (see section above).

In contrast, additional warning information requires additional cognitive processing by the driver. This contradicts vehicle warning guidelines that demand that "a driver should not be required to transpose, compute, interpolate, or translate displayed crash avoidance warning information" (COMSIS Corporation, 1996) because this could delay drivers' responses. However, these costs can be influenced by warning design, and we propose to minimize them by an optimized coding of information.

Presenting natural sounds or shapes that are highly familiar to drivers is an effective way of coding information about the type of hazard. Auditory icons that imitate real-world events (Gaver, 1986) can inherently convey the cause of a warning. Using this type of information presentation, Graham (1999) reported faster but less accurate responses, and McKeown and Isherwood (2007) reported faster and even more accurate responses to respective automotive collision warnings compared to abstract tones. Nakata et al. (2002) measured higher acceptance of visual vehicle collision warnings with specific icons compared to ones with general icons.

Zarife (2014) compared early warnings containing visual information about the type or the location of hazards in various traffic scenarios. While the object cues showed only a few effects, the directional cues clearly improved gaze reactions, and braking responses as well as collision frequencies. A benefit of spatial visual information on subjective evaluation of early warnings has also been reported by Naujoks and Neukum (2014a). In another experiment, verbal information about the direction of cross traffic running a red light led to quicker braking responses, more adapted deceleration, and better subjective ratings (Zhang et al., 2015). Spatially presented warning tones led to faster gaze alignment towards lateral hazards and increased head rotations after false alarms (Zarife, 2014). Further related findings include faster hazard detections and driving reactions in various traffic scenarios (Ho and Spence, 2005; Ho et al., 2006) and shorter stimulus-response times in studies from cognitive psychology (Posner and Boies, 1971; Posner et al., 1980). Nevertheless, in a study by Yan et al. (2014), spatial warning sounds caused no benefit for early warnings and even more collisions for late warnings.

Furthermore, specific warning information could support the comprehension of the cause of a false alarm and, thus, counteract the potential loss of trust and compliance. The theoretical reasoning is basically the same that we outlined earlier with respect to sight obstructions, just vice versa. A false alarm that is incomprehensible for a driver because the initial cause is hidden could become comprehensible with all the relevant information being conveyed by the warning. Accordingly, Lee and Patterson (1993) observed higher subjective reliability for auditory cockpit alarms that contained spatial information. Entin et al. (1996) measured higher trust of operators in visual automatic target detection systems after an explanation of the target selection has been shown. Only few studies addressed specific warning information about occluded hazards. Lee et al. (2002) reported that drivers ignored warnings more often when they were not able to perceive their causes. Thoma et al. (2008) also assumed (but did not prove) that specific icons are more beneficial when the reason for the warning is not visible to the driver. Therefore, further insights are necessary to provide a scientific foundation for the design of future warnings.

1.3. Modality of warnings

The related work described above showed that auditory as well as visual specific warning information can improve the effectiveness of warnings. However, comparing the results of the studies, there are negative as well as positive results for both modalities.

Theories of attention and information processing provide arguments for both modalities, too.

As driving primarily occupies visual resources, models of limited attentional resources (Wickens, 2008) suggest that auditory messages cause less interference with the driving task and less mental overload. However, as auditory warnings are omnidirectional and hard to ignore, they are more prone to cause annoyance (Marshall et al., 2007) which could impair compliance especially after false alarms. Regarding processing of specific information, visual presentation of warnings, such as icons, is more effective than auditory alternatives (Cao et al., 2009).

Many related studies also examined multimodal warnings that have proved to have several advantages for real-world applications (e.g. Ho et al., 2007). However, for research purposes, findings about the individual effects of the involved modalities would be important, but have been confounded at corresponding studies. Taken together, it is still unclear which of the two modalities is more effective in transmitting specific warning information about locations and types of occluded hazards.

1.4. Augmented reality warnings

New display technologies like three-dimensional (3D) audio and augmented reality (AR) displays enable "an integration of synthetic information into the real environment" (Bimber and Raskar, 2005). With AR warnings, the costs of cognitive processing of spatial information should be minimal because the drivers' attention is directly guided towards the hazard without looking at a separate display, and without the need of mentally transposing an abstract representation of the spatial information. There are several studies that support this assumption.

Chen et al. (2007) and Fagerlönn (2011) examined spatial auditory icons conveying type and location of hazards to support drivers' situation awareness in a driving simulator. Compared to arbitrary sounds, they observed shorter learning times, better response performance, and higher satisfaction of the subjects.

In a simulator study by Plavsic et al. (2009), spatially registered symbols that were virtually attached to hazardous vehicles caused fewer collisions and better subjective ratings compared to unregistered warnings.

Schall et al. (2012) and Rusch et al. (2013) examined spatially referenced AR frames virtually attached around relevant road signs, pedestrians, and cars. The AR cues improved hazard detection and compromised neither the detection of irrelevant secondary objects nor the maintenance of the distance to a car in front.

However, the two studies cited above as well as further related work compared AR warnings to a baseline condition without warnings instead of non-AR warnings (e.g. McDonald, 2016). These experimental designs are not suitable to examine the impact of AR coding in terms of warning design. In addition, many studies about AR warnings had limited practical relevance for our application due to divergent technological assumptions like extremely early warning output (e.g. Rusch et al., 2013; Werneke and Vollrath, 2013), methodological constraints like the lack of driving data (Fagerlönn, 2011; Schall et al., 2012), or their orientation towards theoretical paradigms (e.g. Lee et al., 2009). Finally, there is no study that systematically compared the effects of AR warnings of different modalities against their unspecific counterparts and the effects after true as well as false alarms.

1.5. Research needs and scope of the study

Regarding the presented empirical findings of specificity, modality, and AR warnings, it makes sense to further examine the impact of these factors on driving behavior with regard to occluded hazards and frequent false alarms. As the relevant technologies are close to being marketable, the combination of design factors and frame conditions has a high practical relevance. The related work points towards a positive impact of specificity on warning effectiveness for both modalities. Therefore, the aim of our study is a detailed analysis of potential costs and benefits of specific auditory and visual AR warnings.

Our main hypothesis is that the higher costs for processing the additional warning information (implying longer processing and later reaction) are outweighed by the benefits. In detail, we expect faster detection of the occluded hazards due to attention guidance and faster and stronger braking reactions due to improved situation awareness, resulting in a higher rate of accident mitigation after the necessary warnings. Moreover, we expect less annoyance (due to a better comprehension of their cause), higher acceptance, and compliance despite false alarms.

2. Method

2.1. Participants

The sample consisted of 88 participants (18 women, 70 men) with a mean age of 31 (range 20–54, SD = 8.56) and an average driving experience of 13.42 years. All participants reported normal hearing and normal or corrected-to-normal vision and had a valid German driver's license. All participants worked for BMW AG but people with a working background in user interfaces or driver assistance systems were excluded. Seven participants were excluded due to aborted sessions after they reported symptoms of simulator sickness or errors in data recording. The 81 participants who completed both parts of the study were distributed over the four experimental conditions (Fig. 1) with groups of 19, 21, 21, and 20 participants.

2.2. Apparatus and stimuli

The test took place in a driving simulator of BMW AG. The fixed-base simulator included a replica of a BMW 5 series cabin, all relevant driving components of a cockpit (for example, steering wheel, pedals, seat, displays) and a 220° and 60 Hz projection for the simulation of the environment. The rearward scene was shown on three plasma screens located behind the replica. The driving simulation included sounds of the experimental vehicle as well as nearby vehicles. The sound simulation was presented via a center speaker positioned under the driver's seat. The volume level inside the cabin while driving at 50 km/h was 50.5 db(A).

The auditory warnings were presented on a 5.1 surround sound system by Bose with a peak level of 66.9 dB(A) for the warning gong and the 3D warning sounds. The AR warnings were displayed on a combiner-screen on the hood. The focus distance of the real environment as well as the AR warnings in a real AR head-up-display (HUD) would be further than the simulated environment and the AR warnings in our study. However, we adjusted the focus distance of the AR HUD to the projection screen of the simulation to omit unwanted binocular depth cues of the warnings that would not be perceivable in a real-world setting. This is because environmental cues as well as AR warnings would lie beyond the maximum of accommodation and vergence at about 6 m (Proffitt and Caudek, 2003). Therefore, the results should not be affected by this aspect of the simulation. Symbols were displayed in monochrome orange (as a red color appeared too urgent in our scenarios) in a size easily legible for all participants.

For the measurement of the gaze and head movement of the subjects, a SmartEye Pro remote eye-tracking system with five cameras was used. The system allows a real-time measurement of gaze and head direction with 50 Hz, synchronized to the 50 Hz data record-

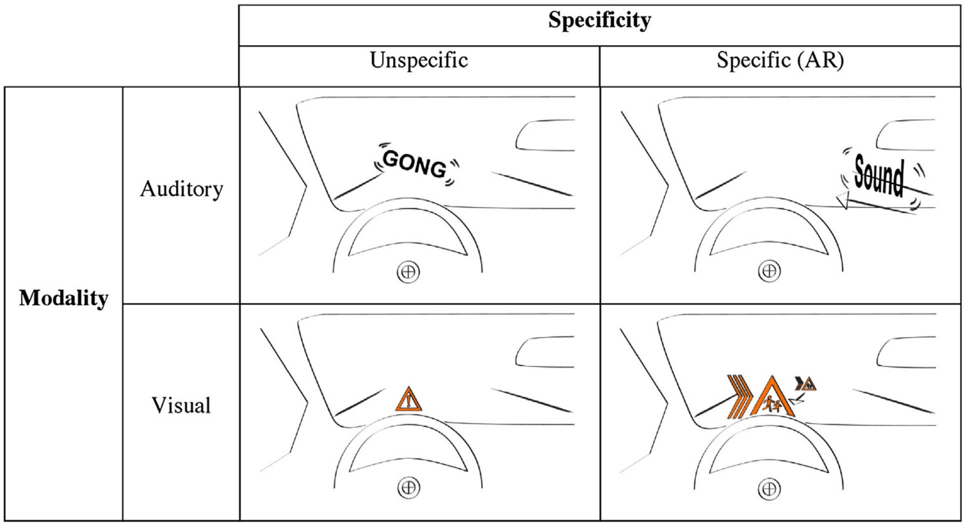

Fig. 1. Schematic depiction of the four examined warning concepts with different modalities and levels of specificity.

Table 1
Sequence of warning situations over part 1 and 2 of the study.

Part	Type of alarm	No.	Scenario
1	True alarms	1.1	Cyclist crossing parallel
		1.2	Pedestrian running on road
		1.3	Cross traffic taking right of way
2	False alarms	2.1	Pedestrian running on road
		2.2	Cyclist crossing parallel
		2.3	Pedestrian running on road
		2.4	Cross traffic taking right of way
		2.5	Pedestrian running on road
		2.6	Cyclist crossing parallel
		2.7	Cross traffic taking right of way
		2.8	Pedestrian running on road

ing of the driving simulator. Tracking availability and quality were observed and rated by the experimenter at the end of each drive as usable or not, and was potentially excluded based on that evaluation.

Within the driving simulation, the participants drove on a suburban course with a speed limit of 50 km/h, medium traffic, and other vehicles parked irregularly on both sides of the road. The route contained 20 intersections with turned-off traffic lights and signs showing the right of way for the participants in all critical scenarios. At four intersections, the participants had to turn right. To exclude possible interferences with the warnings, the turning maneuvers were announced by verbal navigation prompts about 8 s in advance.

2.3. Procedure

Initially, the participants passed a 5-min training to become familiar with the driving simulation. Within the next 5 min, the eye-tracking system was calibrated individually. Then the respective warning concept was introduced to each participant. In the case of the 3D sound warnings, we verified that each participant could distinguish the different types of traffic opponents as well as the different directions. After the preparations, the main test comprised two drives of approximately 10 min each, with a break to answer a questionnaire and to prevent phenomena such as fatigue or simulator sickness from affecting the driving performance.

The first drive included three warning situations (Table 1). All warnings were triggered by the driving simulation when the participant's car exceeded a predefined time-to-arrival (TTA) threshold of 2.87 s to the intersection point of his lane and the trajectory of the potential collision opponent. Previous studies of the same research program determined this timing as appropriate for our type of warnings within similar scenarios (Naujoks and Neukum, 2014b; Seeliger et al., 2014). The hazardous traffic opponents were "waiting" directly behind the edge of the corresponding sight obstruction. They started moving abruptly right before they would have become visible to the participant which corresponded to a TTA of less than 1 s. Because a driver's reaction time plus a maximum deceleration from 50 km/h takes about 1.87 s (Naujoks and Neukum, 2014b), collisions were practically unavoidable without a warning, and we confirmed this assumption in a pre-test with five participants in which none of them could avoid any of the collisions without warnings.

To retain a theoretical benefit of specificity after the first warning, we needed to implement different warning scenarios. Obviously, repeating the same scenario several times would eliminate the potential effect of specificity, as the participants would remember the type and the position of a recurring hazard. To maximize practical relevance, we selected frequent representatives of urban crashes from the GIDAS (German In-Depth Accident Study) database (Technische Universität Dresden, 2013). Based on that, we chose three different crash scenarios that had to incorporate sight obstructions as well as different opponents from different directions.

The first one was a right-turn scenario with a cyclist driving in parallel on the sidewalk, being occluded by billboards until just before the turning and crossing the intersection with constant direction. As the cyclist appeared to the right of the participant after the warning, this was the only scenario where we could measure clearly distinguishable head and gaze reactions, as the cyclist appeared to the right of the participant after the warning. However, because many participants had already touched the brake pedal to prepare for the turn before the warning onset, we could not analyze braking reaction times here. The second scenario included a pedestrian on the right sidewalk, being occluded by a parked van until the participant's car was very close to the pedestrian. The pedestrian ran into the road but changed his direction at the very last moment along the driving path of the participant and continued next to the parked vehicles. We chose this rather unnatural behavior for ethical reasons to make it less distressing for the participants than if they actually hit the pedestrian. Therefore, we could not measure actual collision rates here, but as a surrogate, we analyzed the participants passing speed close to the pedestrian. The third scenario took place at an intersection with cross traffic from the left that violated the participant's right of way. The opponent car was occluded through buildings until the participant's car was very close to the intersection.

The second drive contained eight warning situations that were very similar to the ones in the first except with the variation that they were all false alarms (this fact was not announced to the participants). Additionally, we varied the side of the street or intersection on which the pedestrians and cross traffic appeared (see Table 1). The temporal configurations (for example, timing of warnings) were initially the same as in the first part of our experiment. However, the opponents stopped at the edge of the participants driving path or did not appear at all in the second part.

2.4. Experimental design

The study was designed as a between-subject design with warning modality and specificity as independent variables. Every participant experienced one of four warning concepts representing the four possible combinations with two auditory and two visual warnings with either no additional information or information about type and location of the potential collision opponent (Fig. 1). The participants were assigned randomly to the groups.

Because of the early timing and the expectedly high frequency of false alarms, the warnings were designed to be rather unobtrusive but informative. The unspecific auditory warning was a standard information gong which is already used in current BMW cars, for such things as ice warnings. We decided to employ this sound, as alternative collision warning sounds appeared too urgent in our scenarios. The specific auditory warning was a surround-sound recording of the respective road user. We recorded sounds of a car, a bicycle (idle motion), and a pedestrian (footsteps) and simulated their 3D movement with specialized software (Longcat AudioStage). The unspecific visual warning was an amber warning sign, displayed on the HUD. The specific visual warning contained the same outline of the unspecific warning sign, but instead of the exclamation mark, they contained a symbol for each type of collision opponent (Fig. 2). To control for potential effects of display position, we matched the initial size as well as the vertical position of the unspecific and the specific (AR) visual warning. Besides that, we added arrows that depicted the moving direction of the potential collision opponent. For the visual AR warnings, the sign was rendered by the simulation spatially referenced at the point of intersection of the opponents moving paths.

In order to strictly follow the principles of AR, the sign would have had to be attached directly to the potential collision opponents. However, there was no realistic integrated technical solution available for that. To evaluate a technically realistic implementation (based on the proposals of technical experts), we developed AR warnings that could be displayed on a combiner-screen with $22^\circ \times 9^\circ$ in front of the driver.

As dependent variables, we analyzed gaze and driving data as well as subjective measures. Gaze direction was categorized into four areas of interest: instrument cluster, road ahead, right-hand, and left-hand side of the road. Fixation time of the instrument cluster was accumulated within the time frame between warning onset and passing the potential collision opponent. Time to fixation was assessed as the duration between the warning onset and the driver looking into the direction of the hazard after its appearance, defined by the gaze vector hitting the predefined angular range in which the hazard appeared. Braking response time was calculated as the duration between the onset of the warning and the time when the brake-pedal pressure exceeded 2% of its maximum value. The maximum brake-pedal pressure was compiled within a time frame of 3 s after the warning. Passing speed was the participant's speed at the intersection point of his lane and the trajectory of the potential collision opponent. Finally, collision frequencies were observed in the first and the third scenario.

After each warning, the researcher asked the participant to verbally rate the annoyance of the warning ("How annoying would you rate this warning?") on a 7-point Likert-type rating scale from "not at all" to "very much". After each drive, the participants answered an AttrakDiff questionnaire (Hassenzahl, 2004) as a measure of acceptance of the warning system. The survey provides differential measures of pragmatic quality, which is defined as the effectiveness to achieve behavioral goals (avoiding accidents in our case), and hedonic quality, which aggregates items like "aesthetic" or "creative". Therefore, it does not measure a single score of acceptance but subsumes factors that should determine users' willingness to use a system in terms of acceptance. The AttrakDiff uses items that are very similar to the acceptance scale from Van der Laan et al. (1997).

2.5. Statistical analysis

For some of the dependent variables there were missing data for technical (e.g. low quality of eye-tracking data) or logical reasons (e.g. missing braking reaction time for participants who did not brake at all after warnings). In particular, the quality of gaze tracking data was sufficient for just 45 of the 81 participants. However, this data loss was distributed randomly over the four experimental groups, resulting in remaining cell sizes of 10, 14, 10 and 11 subjects. Furthermore, there were a few questionnaires that were incomplete, and a storage failure in the driving data of one subject in part 2 of the study.

The results of part one of our study were analyzed individually for each warning scenario. For the second part, we summarized the measures as mean values for every participant over the seven situations. For the continuous dependent variables, we conducted two-way factorial 2×2 ANOVAs with modality (auditory, visual) and specificity (unspecific, specific) as between-subject factors. In a few cases, assumptions underlying parametric tests (normal distribution and homogeneity of variances) were violated following the Kolmogorov–Smirnov and Levene test. For a consistent testing procedure over the dependent measures, we ignored these violations, as ANOVA is generally robust despite this (Bortz, 2005). In some cases of significant interaction effects, we conducted planned contrasts with Bonferroni corrected *t*-tests to clarify the impact of specificity for each modality individually. For the observed collisions as a dichotomous variable, we calculated logistic regressions with the same factors as the ANOVAs. All tests were conducted as two-sided with a significance criterion of alpha = 0.05.

3. Results

3.1. Part 1: necessary warnings

Statistical results are summarized in Table A1 in the Appendix A.

3.1.1. Scenario 1: crossing cyclist

In the right-turn scenario with the crossing bicyclist, there was no effect of modality for any of the three observed variables (Fig. 3). Time to fixation, $F(1,41) = 0.92$, $p = 0.343$, $\eta_p{}^2 = 0.02$, maximum brake-pedal pressure, $F(1,77) = 0.55$, $p = 0.459$, $\eta_p{}^2 = 0.01$, and frequency of collisions, $\chi^2 = 0.05$, df = 1, $p = 0.816$, $OR = 0.75$, did not differ between auditory and visual warnings.

Specificity had a significant main effect on time to fixation, $F(1,41) = 8.13$, $p = 0.007$, $\eta_p{}^2 = 0.16$, maximum brake-pedal pressure, $F(1,77) = 6.19$, $p = 0.015$, $\eta_p{}^2 = 0.07$, and collisions, $\chi^2 = 11.42$, df = 1, $p < 0.001$, $OR = 0.14$. With specific warnings, the drivers looked more

Fig. 2. Specific visual AR warning in a scenario with cross traffic taking right of way from the subject's point of view.

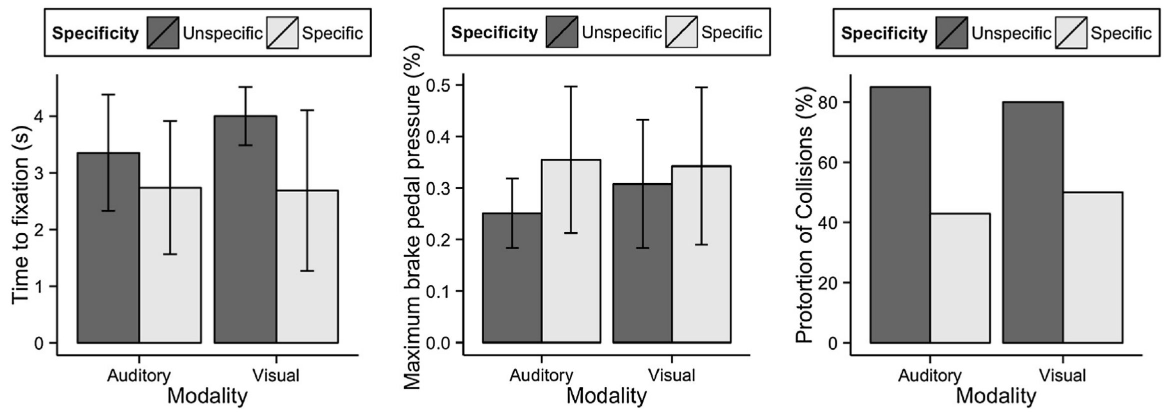

Fig. 3. Mean values of time to fixation, maximum brake-pedal pressure, and proportion of collisions for the different warning concepts. Error bars represent ±1 *SD*.

quickly in the right direction, braked more firmly, and could mitigate more collisions.

There was no significant interaction between modality and specificity either for time to fixation, (1.41) = 1.10, $p = 0.300$, $\eta_p{}^2 = 0.02$ or for maximum brake-pedal pressure, $F(1,77) = 1.52$, $p = 0.221$, $\eta_p{}^2 = 0.02$, or collision frequencies, $\chi^2 = 0.30$, df = 1, $p = 0.583$, $OR = 1.78$.

3.1.2. Scenario 2: crossing pedestrian

In the second scenario with the hidden pedestrian (Fig. 4), modality had a significant main effect on braking reaction time, $F(1,77) = 28.29$, $p < 0.001$, $\eta_p{}^2 = 0.24$, a close to significant main effect on maximum brake-pedal pressure, $F(1,77) = 3.04$, $p = 0.085$, $\eta_p{}^2 = 0.04$, and a significant main effect on passing speed, $F(1,77) = 24.75$, $p < 0.001$, $\eta_p{}^2 = 0.21$.

Specificity had no significant main effect on braking reaction time, $F(1,77) = 1.28$, $p = 0.261$, $\eta_p{}^2 = 0.01$, or on maximum brake-pedal pressure, $F(1,77) = 0.54$, $p = 0.466$, $\eta_p{}^2 = 0.01$, but had a significant main effect on passing speed, $F(1,77) = 6.8$, $p = 0.011$, $\eta_p{}^2 = 0.06$.

The interaction between modality and specificity was significant or close to significant for braking reaction time, $F(1,77) = 11.83$, $p = 0.001$, $\eta_p{}^2 = 0.10$, maximum brake-pedal pressure, $F(1,77) = 3.62$, $p = 0.061$, $\eta_p{}^2 = 0.04$, and passing speed, $F(1,77) = 7.99$, $p = 0.006$, $\eta_p{}^2 = 0.07$.

Contrasts between the auditory concepts did not show significant effects of specificity, either for braking reaction time, $F(1,38) = 3.10$, $p = 0.173$, $\eta_p{}^2 = 0.08$, or for maximum brake-pedal pressure, $F(1,38) = 0.70$, $p = 0.819$, $\eta_p{}^2 = 0.02$, or passing speed, $F(1,38) = 0.40$, $p = 1.000$, $\eta_p{}^2 = 0.00$. The contrasts for the visual concepts confirmed significantly faster braking reactions, $F(1,38) = 9.15$, $p = 0.009$, $\eta_p{}^2 = 0.19$, as well as lower passing speeds, $F(1,38) = 13.91$, $p = 0.001$, $\eta_p{}^2 = 0.27$, after specific warnings. The contrast for specificity of the visual warnings on maximum brake-pedal pressure did not reach significance, $F(1,38) = 3.94$, $p = 0.109$, $\eta_p{}^2 = 0.09$.

3.1.3. Scenario 3: cross traffic

In the third scenario with cross traffic at an intersection, there were only a few effects on objective measures (Fig. 5).

There was no significant main effect of modality on braking reaction time, $F(1,77) = 0.60$, $p = 0.44$, $\eta_p{}^2 = 0.01$, maximum brake-pedal pressure, $F(1,77) = 2.15$, $p = 0.147$, $\eta_p{}^2 = 0.03$, and collisions, $\chi^2 = 1.58$, df = 1, $p = 0.209$ $OR = 0.83$.

Similarly, effects of specificity were far from significant for braking reaction time, $F(1,77) = 1.03$, $p = 0.313$, $\eta_p{}^2 = 0.01$, maximum brake-pedal pressure, $F(1,77) = 2.44$, $p = 0.122$, $\eta_p{}^2 = 0.03$, and collisions, $\chi^2 = 1.04$, df = 1, $p = 0.309$ $OR = 0.94$.

Nonetheless, there was a significant interaction between modality and specificity for braking reaction time, $F(1,77) = 5.84$, $p = 0.018$, $\eta_p{}^2 = 0.07$. Contrasts between the auditory warnings revealed close to significantly slower braking reactions with more specific warnings, $F(1,38) = 5.15$, $p = 0.058$, $\eta_p{}^2 = 0.12$. For the visual warnings, the faster braking reactions after specific warnings did not reach significance, $F(1,38) = 1.31$, $p = 0.519$, $\eta_p{}^2 = 0.03$.

This time the interaction of maximum brake-pedal pressure did not reach significance, nor did the interaction effect of collisions, $\chi^2 = 0.91$, df = 1, $p = 0.341$ $OR = 0.40$, although the absolute collision frequencies with the specific visual warnings were conspicuously lower compared to those shown in Fig. 5.

3.1.4. Overall subjective evaluation

The overall subjective evaluation after the first part of our study revealed a significant main effect of modality for pragmatic qual-

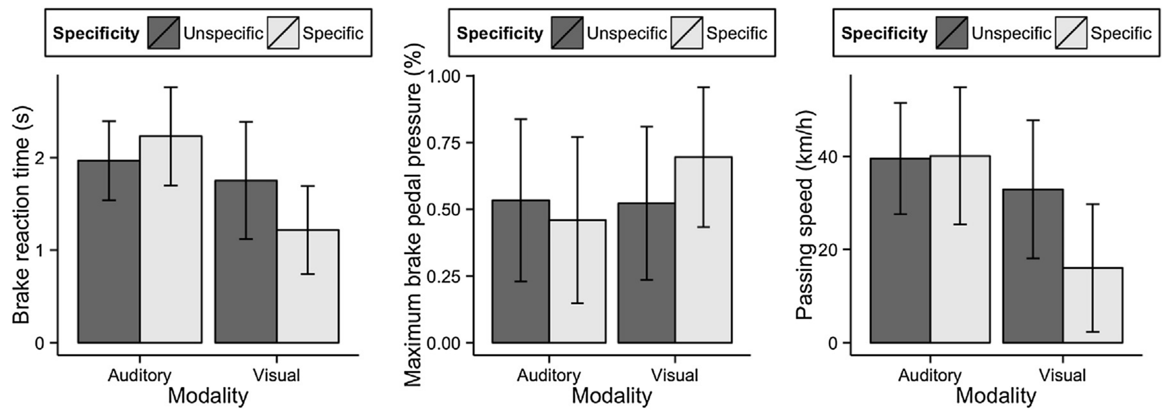

Fig. 4. Mean values of braking reaction time, maximum brake-pedal pressure, and passing speed of the pedestrian for the different warning concepts. Error bars represent ±1 *SD*.

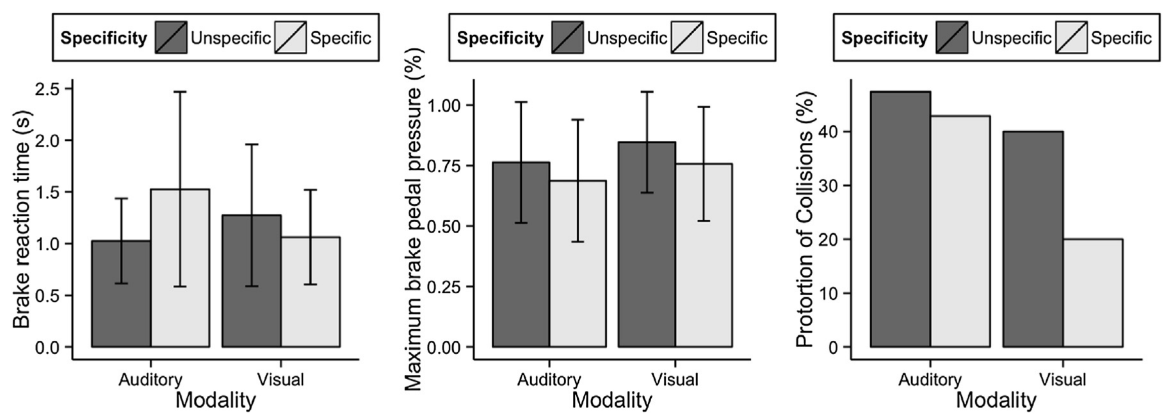

Fig. 5. Mean values of braking reaction time, maximum brake-pedal pressure and number of collisions for the different warning concepts. Error bars represent ±1 *SD*.

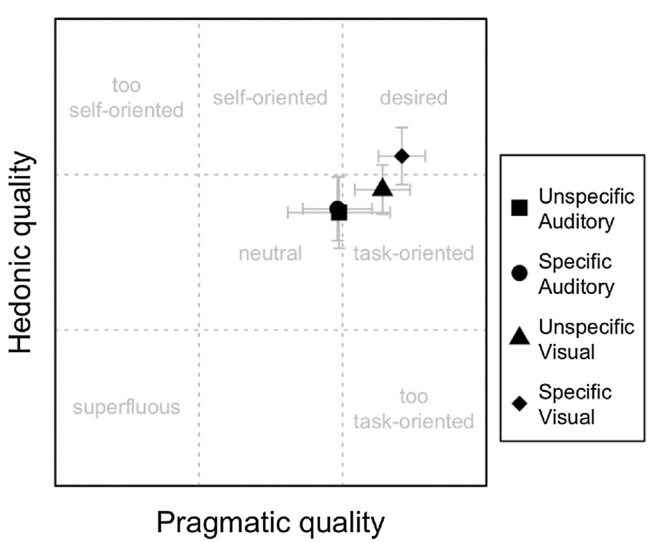

Fig. 6. Portfolio figure of mean pragmatic and hedonic quality of the four concepts after part 1 of the study (Hassenzahl et al., 2008). Error bars represent 95% confidence intervals.

ity, $F(1,75) = 9.94$, $p = 0.002$, $\eta_p{}^2 = 0.12$, as well as hedonic quality, $F(1,75) = 6.84$, $p = 0.011$, $\eta_p{}^2 = 0.08$. The participants rated the visual concepts better than the auditory ones within both categories (Fig. 6).

There was no significant main effect of specificity for the subjective measures of pragmatic quality, $F(1,75) = 0.24$, $p = 0.624$, $\eta_p{}^2 = 0.00$, as well as hedonic quality, $F(1,75) = 1.49$, $p = 0.226$, $\eta_p{}^2 = 0.02$.

Finally, we observed no interaction effects for pragmatic quality, $F(1,75) = 0.37$, $p = 0.548$, $\eta_p{}^2 = 0.01$, as well as hedonic quality, $F(1,75) = 1.03$, $p = 0.314$, $\eta_p{}^2 = 0.01$.

3.2. Part 2: false alarms

Statistical results are summarized in Table A2 in the Appendix A. The pattern of effects of braking reaction time remained present in part two of the study (Fig. 7). Again, there was a significant main effect of modality, $F(1,76) = 5.17$, $p = 0.026$, $\eta_p{}^2 = 0.06$, no significant main effect of specificity, $F(1,76) = 0.03$, $p = 0.872$, $\eta_p{}^2 = 0.00$, but a significant interaction between both of them, $F(1,76) = 4.82$, $p = 0.031$, $\eta_p{}^2 = 0.06$. While auditory and visual warnings resulted in similar reaction times for the unspecific warnings, specificity slightly prolonged the reaction times in the auditory domain and accelerated the reaction times in the visual domain. However, the contrasts did not reach significance, either between the auditory concepts, $F(1,36) = 0.17$, $p = 1.000$, $\eta_p{}^2 = 0.00$, or between the visual ones, $F(1,36) = 0.04$, $p = 1.000$, $\eta_p{}^2 = 0.00$.

A slightly different pattern emerged for the average passing speeds. There was a significant main effect of modality, $F(1,76) = 5.91$, $p = 0.017$, $\eta_p{}^2 = 0.07$, but no main effect of specificity, $F(1,76) = 1.40$, $p = 0.241$, $\eta_p{}^2 = 0.02$. The interaction did not reach significance. $F(1,76) = 1.33$, $p = 0.252$, $\eta_p{}^2 = 0.02$. This means that the drivers passed the potential collision opponents more slowly after visual warnings.

Subjective annoyance due to unnecessary warnings was affected significantly neither by modality, $F(1,76) = 2.05$, $p = 0.156$,

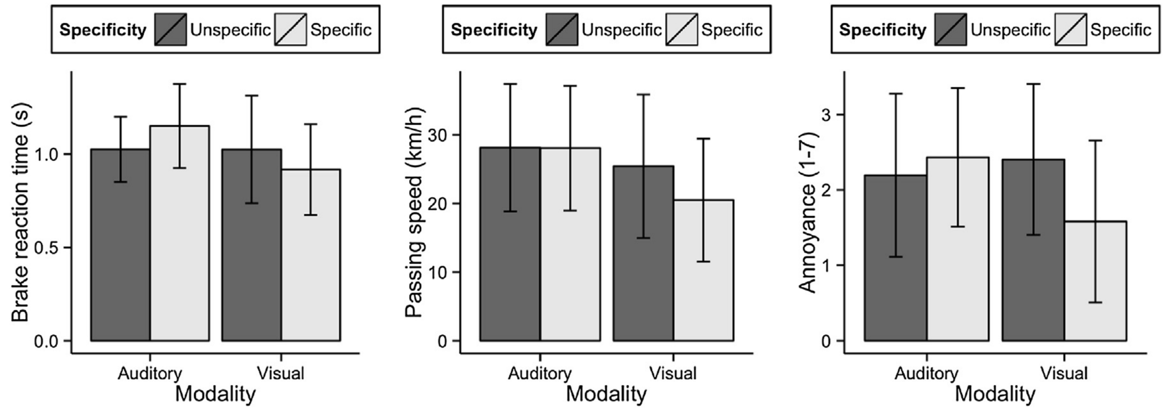

Fig. 7. Mean values of braking reaction time, passing speed, and annoyance rating for the different warning concepts. Error bars represent ±1 *SD*.

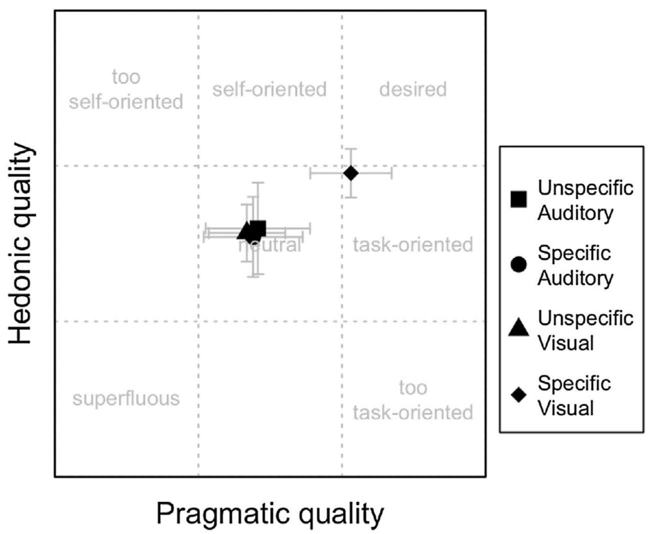

Fig. 8. Portfolio figure of mean pragmatic and hedonic quality of the four concepts after part 2 of the study (Hassenzahl et al., 2008). Error bars represent 95% confidence intervals.

$\eta_p{}^2 = 0.02$, nor by specificity alone, $F(1,76) = 1.65$, $p = 0.203$, $\eta_p{}^2 = 0.02$. Nonetheless, the interaction effect reached significance $F(1,76) = 5.40$, $p = 0.023$, $\eta_p{}^2 = 0.06$. Specificity led to higher annoyance for auditory warnings but to lower annoyance for visual warnings. While the contrast between the auditory concepts did not reach significance, $F(1,36) = 0.32$, $p = 1.000$, $\eta_p{}^2 = 0.01$, specificity significantly reduced annoyance within the visual concepts, $F(1,36) = 7.12$, $p = 0.023$, $\eta_p{}^2 = 0.17$.

Compared to the first part of the study, the pattern of the overall subjective evaluation changed somewhat after the second part of the study.

For pragmatic quality the main effect of modality was close to significant, $F(1,73) = 3.50$, $p = 0.065$, $\eta_p{}^2 = 0.05$, the main effect of specificity significant, $F(1,73) = 5.05$, $p = 0.028$, $\eta_p{}^2 = 0.07$, and the interaction between modality and specificity was also significant, $F(1,73) = 6.26$, $p = 0.015$, $\eta_p{}^2 = 0.08$.

The evaluation of the hedonic quality also revealed a close to significant main effect of modality, $F(1,74) = 2.88$, $p = 0.094$, $\eta_p{}^2 = 0.04$, but no significant main effect of specificity, $F(1,74) = 2.27$, $p = 0.136$, $\eta_p{}^2 = 0.03$. However, the interaction effect again reached significance, $F(1,74) = 4.03$, $p = 0.048$, $\eta_p{}^2 = 0.05$.

The AttrakDiff output in Fig. 8 shows the pattern behind these statistical results with a distinctively better evaluation of pragmatic and hedonic quality of the specific visual warning compared to the other warning concepts that are very close.

4. Discussion

The aim of the study was to examine the effects of modality and specificity on warnings of hidden dangers. Therefore, we analyzed drivers' reactions to four different warning concepts after true and false alarms in various traffic scenarios. Compared to the 100% collisions in a pre-test without warnings, the collision frequencies between 20% and 80% indicate (but do not prove) a general effectiveness of warnings of occluded dangers. Concerning the actual variables of interest, we could observe multiple effects, but they differed considerably between the three scenarios in part one as well as between the true alarms in part one and the false alarms in part two of the experiment.

4.1. Effects of modality

The modality of the warnings revealed few significant effects in our setting. In part one of the study, pragmatic as well as hedonic quality was rated higher for visual warnings than for auditory ones, despite no effect of modality in the objective measures. In part two of the study, we observed lower passing speeds after visual warnings compared to auditory ones. However, the inspection of the results for the individual concepts revealed that both effects are related to advantages of the specific visual warnings, and there are only minor advantages of the unspecific visual warnings compared with the auditory ones. Therefore, we cannot infer a general preference for one of the modalities–especially not for the unspecific concepts. There were only small absolute differences between the results of the auditory and visual unspecific warnings, which we found rather surprising in the light of related work that had proved various advantages of each modality (e.g. Scott and Gray, 2008).

A possible explanation for this discrepancy was the particular setting of our study, where none of the preconditions for possible differential advantages were given. The visual attention of the drivers was not distracted, and the visual resources were not occupied by, for example, complex side tasks as a theoretical basis for benefits of auditory cues (Wickens, 2008). On the other hand, the rather low urgency of the warning gongs possibly mitigated the predisposition of auditory cues to annoy the driver (Marshall et al., 2007). The objective of this experimental setting was to represent a highly common and realistic driving task and to maximize the practical relevance of the results. However, it would be interesting to reassess these results for situations with higher visual workload, as a common and accident-related precondition. The interaction between modality and specificity produced more explicit results that are described in Section 4.3.

4.2. Effects of specificity

The main effects of specificity varied considerably over the course of the study. In the first scenario, we observed a clear advantage of specificity in terms of faster gaze responses toward the hazard as well as harder braking reactions resulting in crash rates reduced by around 50%. This indicates that the participants could process the specific information effectively and utilize it to improve avoidance behavior. The fact that this effect emerged for both modalities confirms the general potential of specificity that has already been shown by several studies for individual modalities (e.g. Naujoks and Neukum, 2014a; Zhang et al., 2015). Two characteristics of this scenario probably contributed to this finding. On the one hand, for the participants it was the first confrontation with the warning and the first hazard scenario. Therefore, they had no previous experience as a reference, but only the warning information as a basis for their reaction. Yet, it is unclear whether this situation was more or less representative of real-world applications, in which drivers would experience repeated warnings but at distinctly longer intervals. On the other hand, the cyclist in this scenario was difficult to detect even after his appearance close to the participant's car. Therefore, the specific warnings had a higher informational value in this scenario compared to the subsequent scenarios where the hazard appeared within the central field of view of the drivers.

In the second scenario, the passing speed was still reduced after specific compared to unspecific warnings. But the drivers did not brake more quickly or more firmly after specific warnings this time. However, a subsequent visual inspection of the time courses of the brake-pedal pressure revealed a confounding of the brake reactions to the warnings and to the appearance of the pedestrian itself. After the specific warnings, the brake-pedal pressure was increased more quickly, while there were higher peak levels after unspecific warnings. This pattern might be indicative of a better situation awareness with specific warnings and a stronger startle reaction at the appearance of the pedestrian after unspecific warnings.

After the second scenario, there were no more general advantages of specificity, which contradicted our hypothesis. Generally, we expected a higher impact of the information itself, particularly on the subjective evaluations after false alarms because the drivers should have had a better understanding of their causes. There are at least two possible explanations for this. Firstly, it seems that the benefits of additional information were considerably dependent on specific design parameters within the two modalities. There were still main effects of specificity after scenario two, but they were always associated with interaction effects, which limits their interpretation. Secondly, the benefit of specificity obviously depended on the ambiguity of the corresponding scenario. Because the participants could increasingly anticipate the properties of the scenarios from their previous experiences, the added value inherently decreased over the course of the study.

Accordingly, the expectation of a general benefit of specificity to maintain higher trust and compliance despite false alarms (Maltz and Shinar, 2004, 2007) was not confirmed. As our assumption was based on conclusions drawn from subjective ratings (Entin et al., 1996; Lee and Patterson, 1993), we should have measured subjective trust in hindsight. We suggest further investigating the potentially different mechanisms of trust and affective constructs such as annoyance. A possible explanation for the absence of the expected preservation of compliance using the specific auditory warnings could be a modality-specific effect of annoyance opposing the expected general effect of specificity on trust. In addition, the analysis of the interaction effect in the next section provides an explanation for the lack of a significant main effect.

4.3. Interaction effects of modality and specificity

Besides the first two scenarios, the interaction between modality and specificity was the prevalent effect over the different scenarios and dependent measures. The pattern of the interaction effect was a negative impact of specificity for auditory warnings but a positive impact of specificity for visual warnings. Concerning braking reaction times, this effect emerged for scenarios two and three and for the unnecessary warnings in part two of the study. Furthermore, we observed the same pattern for maximum brake-pedal pressure and passing speed in scenario two. Finally, the effect recurred in terms of the lower annoyance ratings after unnecessary warnings, and the practical and hedonic quality of the specific visual warnings compared to the other warnings at the end of the study. This effect contradicts our hypothesis of a modality-independent advantage of specific warning information. At the same time, however, it emphasizes the potential of the visual AR warnings.

An explanation of these differential effects could be the diverging effectiveness of information transmission to the driver with the two specific concepts. The observed effects over the various measures point to an impact on several stages of the processing of the warning information (Wogalter, 2006). Firstly, the specific visual warnings were presented directly on the road with a clearly perceivable onset that probably facilitated its detection. The approaching sound of the specific auditory warning caused a fade-in effect that smoothed the acoustic onset and could have delayed or even masked its detection. Secondly, the cognitive processing of the specific information about location and type of the hazards was obviously more effective for the visual cues, which extends related work (e.g. Cao et al., 2010) to AR warnings. The inherent coding of the distance by virtually placing the warning sign directly on the road, the arrows for the coding of the lateral location of the hazard, and the characteristic icons for the type of traffic opponent are all commonly used codes in road traffic that should be highly common and easy to process. In contrast, as vehicle cabins get more and more acoustically isolated, spatial auditory environmental information is hardly perceivable in modern cars, and drivers are not used to considering this kind of information except for a few occasions such as locating emergency vehicles via their sirens. Additionally, for the scenarios with the pedestrian or cross traffic, the hazard position stayed in a narrow sector in front of the driver and the auditory distinction between the left and right side of the road was subtle and probably difficult to make, at least compared to the visual arrow symbols. Finally, the distance of the conflict was coded via the sound level of the auditory specific warnings, and despite the fact that this should be a very natural and learned principle, estimation of distances was clearly more precise for visual cues.

Taken together, these findings demonstrate the potential benefits of visual AR warnings and support the idea that besides the amount of information, the exact coding markedly affects the efficacy of such warnings.

4.4. Limitations

Besides the general limitations of a driving simulation, our study contained a high frequency of warning scenarios. However, one of the most important requirements for the validity of collision-warning research is including situations in which the participants generally do not expect a hazard or not a specific type of hazard. This precondition inherently decreases over the course of several situations. With the implementation of three diverse and complex warning scenarios, we tried to control this as far as possible. As a consequence, we had to measure different dependent variables for each scenario and the complexity of the scenarios caused high variance within our measures. We could not design a higher number of different scenarios for the second part of the study, and some participants probably recognized some of them. Furthermore, we did not apply an alpha level correction to control type I error inflation for the multiple measures and scenarios (as recommended by e.g. Feise, 2002; Rothman, 1990). However, we acknowledge the controversy about this topic, and it should be considered for the generalization of the results.

Another implication of the high number of warning scenarios was the choice of a between-subject design. The analysis of the loss of compliance over several false alarms required a certain number of warning scenarios. Using a within-subject design would have multiplied the duration of the study. On the other hand, the alternative design would have had several advantages. The necessary sample size would have been much smaller than the current one. Further, interindividual variability would have been inherently controlled. Nevertheless, we decided to focus on the first confrontation with one of the warning concepts as we had little experience in the applicability of measurements after higher numbers of repetitions.

Another limitation in relation to the generalization of our results is the high number of design parameters of warnings. The unspecific warnings were adopted from real world applications, but even these differ markedly between vehicle manufacturers and models. For example, warnings on HUDs are only available in premium vehicles although they have been proven to have clear advantages. The number of possible design parameters of more specific warnings is even higher. For example, we decided not to attach the specific visual warning directly to the hazard, which would have been a theoretically more consistent spatial referencing, but instead to the road in front of the vehicle as this was technically much more realistic. This probably facilitated the detection of the warning and certainly affected the results compared to the mentioned variant. Moreover, we controlled the initial size and vertical position of the AR warning to match the non-AR warning in the HUD. However, we still cannot exclude possible effects of the slightly different initial vertical position. Furthermore, there was an inherent confounding of spatial referencing and the necessary scaling animation itself, regardless of the exact geometric execution. Further design parameters that were not covered at all by this experiment that could also affect warning effectiveness or could modulate the reported effects are timing, chosen symbols, sizes, sounds, sound levels, and multimodal combinations.

Finally, the quality of the implementation of the specific warnings was limited. The visual warnings were projected on a combiner-screen with a focus distance of about the screen distance of the driving simulation which could have compromised depth perception compared to a real-world application. For the auditory warnings, we tried to reach the best possible quality, but neither our development tools nor the driving simulation were optimized to examine complex acoustics. The peak levels of the different sounds were balanced subjectively, but we did not consider aspects such as signal to noise ratio. Moreover, the setup of the surround-sound system was very basic. The simulation did not include a replica with a tuned surround-sound system but a basic cabin with subpar auditory characteristics, where we integrated consumer grade surround speakers as a workaround. Therefore, the results for the specific auditory warnings might be different with a better implementation of the same concept.

5. Conclusion

The aim of our simulator study was to investigate the effects of modality and specificity of necessary and unnecessary in-vehicle warnings of occluded hazards. Both independent variables showed effects on warning effectiveness and acceptance, but the effects of specificity clearly depended on the modality of the warning. The specific visual warning showed advantages in most of the measures over all scenarios and for necessary and unnecessary warnings. The specific auditory warning performed similarly in the first scenarios, but after the second one, there proved to be no more benefits and even several detriments. A possible reason for this pattern was the lower effectiveness of coding for the auditory compared to the visual concept and the general degradation of added value due to specificity over the course of the study.

Our study extends our knowledge about the impact of warning design on driver behavior with regard to several aspects. Firstly, concepts of future AR warnings were introduced that stand out due to their technical feasibility concerning the underlying hazard detection as well as display technologies. Moreover, while many related studies about AR warnings focused on one concept or modality (e.g. Rusch et al., 2013; Werneke and Vollrath, 2013), this work was based on a systematic variation of two basic factors of warning design. This approach revealed the importance of the exact coding of information (besides the amount of information and modality), the impact of which probably increases with more specific warnings. Concerning sight obstructions, the experiment extends existing subjective indications about the benefit of additional information for warning effectiveness with objective results. The hypothesis concerning the interplay of warning information and false alarms was the second contextual aspect of particular consideration. Although our general expectation was not confirmed, the outcome suggests a more detailed investigation of the related psychological mechanism in future experiments.

To conclude, because our results revealed a high impact of warning design on collision mitigation, the rather complex AR concepts should be analyzed in more detail, and the corresponding psychological mechanisms warrant more attention in future studies. Within the additional information of the visual AR warnings, several information elements and coding techniques were combined. It is unclear to what extent the advantages of these warnings were due to the spatial referencing or the symbolic information regarding type and moving direction of the hazard. The addition of specific symbols to a warning display would be much easier than the implementation of the AR warnings.

Appendix A.

Table A3

Table A1
Statistical results across dependent measures from part 1.

	Modality	Specificity	Interaction
Scenario 1			
Time to fixation of cyclist	$F(1,41) = 0.92$, $p = 0.343$, $\eta_p^2 = 0.01$	$F(1,41) = 8.13$, $p = 0.007$, $\eta_p^2 = 0.17$	$F(1,41) = 1.1$, $p = 0.300$, $\eta_p^2 = 0.03$
Maximum brake pedal pressure	$F(1,77) = 0.55$, $p = 0.459$, $\eta_p^2 = 0.01$	$F(1,77) = 6.19$, $p = 0.015$, $\eta_p^2 = 0.07$	$F(1,77) = 1.52$, $p = 0.221$, $\eta_p^2 = 0.02$
Collisions	$\chi^2 = 0.05$, df = 1, $p = 0.816$, $OR = 0.75$	$\chi^2 = 11.42$, df = 1, $p < 0.000$, $OR = 0.14$	$\chi^2 = 0.30$, df = 1, $p = 0.583$, $OR = 1.78$
Scenario 2			
Brake reaction time	$F(1,77) = 28.29$, $p < 0.001$, $\eta_p^2 = 0.27$	$F(1,77) = 1.28$, $p = 0.261$, $\eta_p^2 = 0.02$	$F(1,77) = 11.83$, $p = 0.001$, $\eta_p^2 = 0.13$
Maximum brake pedal pressure	$F(1,77) = 3.04$, $p = 0.085$, $\eta_p^2 = 0.04$	$F(1,77) = 0.54$, $p = 0.466$, $\eta_p^2 = 0.01$	$F(1,77) = 3.62$, $p = 0.061$, $\eta_p^2 = 0.04$
Speed at passing pedestrian	$F(1,77) = 24.75$, $p < 0.001$ $\eta_p^2 = 0.25$	$F(1,77) = 6.8$, $p = 0.011$, $\eta_p^2 = 0.08$	$F(1,77) = 7.99$, $p = 0.006$, $\eta_p^2 = 0.09$

Table A1 (*Continued*)

	Modality	Specificity	Interaction
Scenario 3			
Brake reaction time	$F(1,77) = 0.6$, $p = 0.440$, $\eta_p{}^2 = 0.01$	$F(1,77) = 1.03$, $p = 0.313$, $\eta_p{}^2 = 0.01$	$F(1,77) = 5.84$, $p = 0.018$, $\eta_p{}^2 = 0.07$
Maximum brake pedal pressure	$F(1,77) = 2.15$, $p = 0.147$, $\eta_p{}^2 = 0.03$	$F(1,77) = 2.44$, $p = 0.122$, $\eta_p{}^2 = 0.03$	$F(1,77) = 0.02$, $p = 0.895$, $\eta_p{}^2 = 0.00$
Collisions	$\chi^2 = 1.58$, df = 1, $p = 0.209$ $OR = 0.83$	$\chi^2 = 1.04$, df = 1, $p = 0.309$ $OR = 0.94$	$\chi^2 = 0.91$, df = 1, $p = 0.341$ $OR = 0.40$
After scenarios 1–3			
Pragmatic Quality	$F(1,75) = 9.94$, $p = 0.002$, $\eta_p{}^2 = 0.12$	$F(1,75) = 0.24$, $p = 0.624$, $\eta_p{}^2 = 0.00$	$F(1,75) = 0.37$, $p = 0.548$, $\eta_p{}^2 = 0.00$
Hedonic Quality	$F(1,75) = 6.84$, $p = 0.011$, $\eta_p{}^2 = 0.08$	$F(1,75) = 1.49$, $p = 0.226$, $\eta_p{}^2 = 0.02$	$F(1,75) = 1.03$, $p = 0.314$, $\eta_p{}^2 = 0.01$

Table A2
Statistical results across dependent measures from part 2 of the study.

	Modality	Specificity	Interaction
Brake reaction time	$F(1,76) = 5.17$, $p = 0.026$, $\eta_p{}^2 = 0.06$	$F(1,76) = 0.03$, $p = 0.872$, $\eta_p{}^2 = 0$	$F(1,76) = 4.82$, $p = 0.031$, $\eta_p{}^2 = 0.06$
Passing speed	$F(1,76) = 5.91$, $p = 0.017$, $\eta_p{}^2 = 0.07$	$F(1,76) = 1.4$, $p = 0.241$, $\eta_p{}^2 = 0.02$	$F(1,76) = 1.33$, $p = 0.252$, $\eta_p{}^2 = 0.02$
Annoyance	$F(1,76) = 2.05$, $p = 0.156$, $\eta_p{}^2 = 0.03$	$F(1,76) = 1.65$, $p = 0.203$, $\eta_p{}^2 = 0.02$	$F(1,76) = 5.4$, $p = 0.023$, $\eta_p{}^2 = 0.07$
Pragmatic quality	$F(1,73) = 3.5$, $p = 0.065$, $\eta_p{}^2 = 0.05$	$F(1,73) = 5.05$, $p = 0.028$, $\eta_p{}^2 = 0.06$	$F(1,73) = 6.26$, $p = 0.015$, $\eta_p{}^2 = 0.08$
Hedonic quality	$F(1,74) = 2.88$, $p = 0.094$, $\eta_p{}^2 = 0.04$	$F(1,74) = 2.27$, $p = 0.136$, $\eta_p{}^2 = 0.03$	$F(1,74) = 4.03$, $p = 0.048$, $\eta_p{}^2 = 0.05$

Table A3
Cell sizes of warning concepts for each dependent variable.

Scenario or part	Dependent variable	Auditory unspecific	Auditory specific	Visual unspecific	Visual specific
Scenario 1	Maximum brake pedal pressure (%)	19	21	20	20
Scenario 1	Proportion of collisions	19	21	20	20
Scenario 1	Time to fixation (s)	10	14	10	11
Scenario 2	Maximum brake pedal pressure (%)	19	21	20	20
Scenario 2	Brake reaction time (s)	19	21	20	20
Scenario 2	Passing speed (km/h)	19	21	20	20
Scenario 3	Maximum brake pedal pressure (%)	19	21	20	20
Scenario 3	Brake reaction time (s)	19	21	20	20
Scenario 3	Proportion of collisions	19	21	20	20
Part 1	Pragmatic Quality (1–7)	19	21	18	20
Part 1	Hedonic Quality (1–7)	19	21	18	20
Part 2	Brake reaction time (s)	17	13	17	18
Part 2	Passing speed (km/h)	18	20	19	19
Part 2	Annoyance (1–7)	18	20	19	19
Part 2	Pragmatic Quality (1–7)	18	20	19	18
Part 2	Hedonic Quality (1–7)	18	20	19	19

References

Bimber, O., Raskar, R., 2005. Spatial Augmented Reality Merging Real and Virtual Worlds. CRC Press, Wellesley.

Bliss, J.P., Gilson, R.D., Deaton, J.E., 1995. Human probability matching behaviour in response to alarms of varying reliability. Ergonomics 38 (11), 2300–2312.

Bortz, J., 2005. Statistik: Für Human- und Sozialwissenschaftler, 6th ed. Springer Medizin Verlag, Heidelberg.

Breznitz, S., 1984. Cry Wolf: The Psychology of False Alarms. Lawrence Erlbaum Associates, New Jersey.

COMSIS Corporation, 1996. Preliminary Human Factor Guidelines for Crash Avoidance Warning Devices. Silver Spring, MD.

Campbell, J.L., Richard, C.M., Brown, J.L., McCallum, M., 2007. Crash Warning System Interfaces: Human Factors Insight and Lessons Learned (Washington, DC).

Cao, Y., Castronovo, S., Mahr, A., Müller, C., 2009. On timing and modality choice with local danger warnings for drivers. Proceedings of the 1 st International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 75.

Cao, Y., Mahr, A., Castronovo, S., Stahl, C., Christian, M., Theune, M., Müller, C.A., 2010. Local danger warnings for drivers: the effect of modality and level of assistance on driver reaction. In: Proceeding of the 14th International Conference on Intelligent User Interfaces, ACM, pp. 239–248.

Chen, F., Qvint, G., Jarlengrip, J., 2007. Listen! there are other road users close to you ?Improve the traffic awareness of truck drivers. Proceedings of the 4th International Conference on Universal Access in Human-computer Interaction: Ambient Interaction, 323–329.

Chugh, J.S., Caird, J.K., 1999. In-Vehicle train warnings (ITW): the effect of reliability and failure type on driver perception response time and trust. Proc. Hum. Factors Ergon. Soc. Ann. Meeting 43 (18), 1012–1016.

Dillon, A., 2001. User acceptance of information technology. In: Karwowski, W. (Ed.), Encyclopedia of Human Factors and Ergonomics. Taylor and Francis, London.

Endsley, M.R., 1995. Toward a theory of situation awareness in dynamic systems. Hum. Factors: J. Hum. Factors Ergon. Soc. 37 (1), 32–64.

Entin, E.B., Entin, E.E., Serfaty, D., 1996. Optimizing aided target-Recognition performance. Proc. Hum. Factors Ergon. Soc. Ann. Meeting 40 (4), 233–237.

Fagerlönn, J., 2011. Urgent alarms in trucks: effects on annoyance and subsequent driving performance. IET Intel. Transport Syst. 5 (4), 252–258.

Feise, R.J., 2002. Do multiple outcome measures require p-value adjustment? BMC Med. Res. Methodol. 2 (1), 1.

Fuerstenberg, K., Hopstock, M., Obojski, A., Rössler, B., Chen, J., Deutschle, S., et al., 2007. INTERSAFE Deliverable D 40.75, Final Report (Project Evaluation and Effectiveness of the Intersection Safety System), Hamburg. Retrieved from http://www.prevent-ip.org/en/prevent_subprojects/intersection_safety/intersafe/.

Gaver, W., 1986. Auditory icons: using sound in computer interfaces. Hum.-Comput. Interact. 2 (2), 167–177.

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.

Graham, R., 1999. Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application. Ergonomics 42 (9), 1233–1248.

Green, P., Levison, W., Paelke, G., Serafin, C., 1995. Preliminary Human Factors Design Guidelines for Driver Information Systems (McLean, VA).

Hägglund, C., 2008. Design of a Driver-vehicle Interface for Local Surrounding World Information in Intersections Surrounding World Information in Intersections. Uppsala Universitet.

Hassenzahl, M., Burmester, M., Koller, F., 2008. Der User Experience auf der Spur: Zum Einsatz von www.attrakdiff.de. Usability Professionals.

Hassenzahl, M., 2004. The interplay of beauty, goodness, and usability in interactive products. Hum.-Comput. Interact. 19 (4), 319–349.

Ho, C., Spence, C., 2005. Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. J. Exp. Psychol.: Appl. 11 (3), 157–174.

Ho, C., Tan, H.Z., Spence, C., 2006. The differential effect of vibrotactile and auditory cues on visual spatial attention. Ergonomics 49 (7), 724–738.

Ho, C., Reed, N., Spence, C., 2007. Multisensory In-Car warning signals for collision avoidance. Hum. Factors: J. Hum. Factors Ergon. Soc. 49 (6), 1107–1114.

Informal Group on Intelligent, & Transport Systems, 2011. Guidelines on Establishing Requirements for High-priority Warning Signals. Informal Group on Intelligent, & Transport Systems, Geneva.

Kiefer, R.J., LeBlanc, D., Palmer, M., Salinger, J., Deering, R., Shulman, M., 1999. Development and Validation of Functional Definitions and Evaluation Procedures For Collision Warning/Avoidance Systems (Washington, DC).
LeBlanc, D.J., Tsimhoni, O., 2008. Providing Drivers With Road-Edge Information to Reduce Road Departure Crashes in a Military Vehicle Fleet, Vol. 7 (Ann Arbor, MI).
Lee, M.D., Patterson, R.W., 1993. The application of three-dimensional audio displays to aircraft cockpits: user requirements, technology assessment, and operational recommendations. In: Proceedings of the Beyond Speech/Virtual Reality/Teleoperation, New York, NY : SIG Advanced Applications.
Lee, J.D., See, K.A., 2004. Trust in automation: designing for appropriate reliance. Hum. Factors: J. Hum. Factors Ergon. Soc. 46 (1), 50–80.
Lee, J.D., McGehee, D.V., Brown, T.L., Reyes, M.L., 2002. Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. Hum. Factors: J. Hum. Factors Ergon. Soc. 44 (2), 314–334.
Lee, Y.-C., Lee, J.D., Ng Boyle, L., 2009. The interaction of cognitive load and attention-Directing cues in driving. Hum. Factors: J. Hum. Factors Ergon. Soc. 51 (3), 271–280.
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.
Lees, M.N., 2010. Context Conditions Drivers' Disposition Towards Alarms. University of Iowa (Retrieved from) http://ir.uiowa.edu/etd/845.
Lerner, N.D., Dekker, D.K., Steinberg, G.V., Huey, R.W., 1996. Inappropriate Alarm Rates and Driver Annoyance (Washington, DC).
Lu, M., Wevers, K., Van Der Heijden, R., 2005. Technical feasibility of advanced driver assistance systems (ADAS) for road traffic safety. Transp. Plann. Technol. 28 (3), 167–187.
Maltz, M., Shinar, D., 2004. Imperfect in-vehicle collision avoidance warning systems can aid drivers. Hum. Factors: J. Hum. Factors Ergon. Soc. 46 (2), 357–366.
Maltz, M., Shinar, D., 2007. Imperfect in-vehicle collision avoidance warning systems can aid distracted drivers. Transp. Res. Part F: Traffic Psychol. Behav. 10 (4), 345–357.
Marshall, D.C., Lee, J.D., Austria, P.A., 2007. Alerts for In-Vehicle information systems: annoyance, urgency, and appropriateness. Hum. Factors: J. Hum. Factors Ergon. Soc. 49 (1), 145–157.
McDonald, Z., 2016. Augmented Reality HUDs: Warning Signs and Drivers' Situation Awareness. Rochester Institute of Technology.
McKeown, D., Isherwood, S., 2007. Mapping candidate within-Vehicle auditory displays to their referents. Hum. Factors: J. Hum. Factors Ergon. Soc. 49 (3), 417–428.
Nakata, A., Campbell, J.L., Richman, J.B., 2002. Driver acceptance of general vs. specific icons for In-Vehicle information. Proc. Hum. Factors Ergon. Soc. Ann. Meeting 46 (22), 1799–1803.
Naujoks, F., Neukum, A., 2014a. Specificity and timing of advisory warnings based on cooperative perception. In: Butz, A., Koch, M., Schlichter, J. (Eds.), Mensch & Computer 2014 Workshopband. De Gruyter Oldenbourg, Berlin, pp. 229–238.
Naujoks, F., Neukum, A., 2014b. Timing of in-vehicle advisory warnings based on cooperative perception. In: de Waard, D., Brookhuis, K., Wiczorek, R., di Nocera, F., Brouwer, R., Barham, P., Weikert, A., Kluge, W., Gerbino, A. (Eds.), Proceedings of the Human Factors and Ergonomics Society Europe Chapter Annual Meeting. , pp. 193–206.
Naujoks, F., Grattenthaler, H., Neukum, A., Weidl, G., Petrich, D., 2014. Effectiveness of advisory warnings based on cooperative perception. IEEE Intell. Vehicles Symp. Proceed., 246–252.
Parasuraman, R., Riley, V., 1997. Humans and automation: use, misuse, disuse, abuse. Hum. Factors: J. Hum. Factors Ergon. Soc. 39 (2), 230–253.
Parasuraman, R., Hancock, P.A., Olofinboba, O., 1997. Alarm effectiveness in driver-centred collision-warning systems. Ergonomics 40 (3), 390–399.
Pierowicz, J., Jocoy, E., Lloyd, M., Bittner, A., Pirson, B., 2000. Intersection Collision Avoidance Using ITS Countermeasures, Final Report: Performance Guidelines (Washington, DC).
Plavsic, M., Bubb, H., Duschl, M., Tönnis, M., Klinker, G., 2009. Ergonomic design and evaluation of augmented reality based cautionary warnings for driving assistance in urban environments. In: Proceedings of the 17th World Congress on Ergonomics, Bejing China: International Ergonomics Association.
Posner, M.I., Boies, S.J., 1971. Components of attention. Psychol. Rev. 78 (5), 391–408.
Posner, M.I., Snyder, C.R., Davidson, B.J., 1980. Attention and the detection of signals. J. Exp. Psychol. 109 (2), 160–174.
Proffitt, D.R., Caudek, C., 2003. Depth perception and the perception of events. In: Healy, A.F., Proctor, R.W. (Eds.), Handbook of Psychology: Experimental Psychology. , 4th ed. John Wiley & Sons, Hoboken, New Jersey.
Rothman, K.J., 1990. No adjustments are needed for multiple comparisons. Epidemiology 1 (1), 43–46.
Rusch, M.L., Schall, M.C., Gavin, P., Lee, J.D., Dawson, J.D., Vecera, S., Rizzo, M., 2013. Directing driver attention with augmented reality cues. Transp. Res. Part F: Traffic Psychol. Behav. 16 (319), 127–137.
Schall, M.C., Rusch, M.L., Lee, J.D., Dawson, J.D., Thomas, G., Aksan, N., Rizzo, M., 2012. Augmented reality cues and elderly driver hazard perception. Hum. Factors: J. Hum. Factors Ergon. Soc. 55 (3), 643–658.
Scott, J.J., Gray, R., 2008. A comparison of tactile, visual, and auditory warnings for rear-End collision prevention in simulated driving. Hum. Factors: J. Hum. Factors Ergon. Soc. 50 (2), 264–275.
Seeliger, F., Weidl, G., Petrich, D., Naujoks, F., Breuel, G., Neukum, A., Dietmayer, K., 2014. Advisory warnings based on cooperative perception. In: IEEE Intelligent Vehicles Symposium, Proceedings. IEEE, Dearborn, Michigan, pp. 246–252.
Sorkin, R.D., Kantowitz, B.H., Kantowitz, S.C., 1988. Likelihood alarm displays. Hum. Factors: J. Hum. Factors Ergon. Soc. 30 (4), 445–459.
Staubach, M., 2009. Factors correlated with traffic accidents as a basis for evaluating Advanced Driver Assistance Systems. Accid. Anal. Prevent. 41 (5), 1025–1033.
Technische Universität Dresden, 2013. GIDAS – German In-Depth Accident Study (Retrieved November 21, 2013, from) http://www.vufo.de/forschung-und-entwicklung/gidas/.
Thoma, S., Lindberg, T., Klinker, G., 2008. Evaluation of a generic warning for multiple intersection assistance systems. In: de Waard, D., Godthelp, H., Kooi, F., Brookhuis, K. (Eds.), Human Factors, Security and Safety. Shaker Publishing, Maastricht, pp. 173–188.
Van der Laan, J.D., Heino, A., De Waard, D., 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. Transp. Res. Part C: Emerg. Technol. 5 (1), 1–10.
Werneke, J., Vollrath, M., 2013. How to present collision warnings at intersections?: —A comparison of different approaches. Accid. Anal. Prevent. 52, 91–99.
Wickens, C.D., 2008. Multiple resources and mental workload. Hum. Factors: J. Hum. Factors Ergon. Soc. 50 (3), 449–455.
Wogalter, M.S., 2006. Communication-Human information processing (C-HIP) model. In: Wogalter, M.S. (Ed.), Handbook of Warnings. CRC Press, Mahwah, NJ, pp. 51–61.
Yan, X., Xue, Q., Ma, L., Xu, Y., 2014. Driving-simulator-based test on the effectiveness of auditory red-light running vehicle warning system based on time-to-collision sensor. Sensors (Switzerland) 14 (2), 3631–3651.
Zabyshny, A., Ragland, D., 2003. False Alarms and Human-Machine Warning Systems (Berkley).
Zarife, R., 2014. Integrative Warning Concept for Multiple Driver Assistance Systems. Julius-Maximilians-Universität Würzburg, Würzburg.
Zhang, Y., Yan, X., Yang, Z., 2015. Discrimination of effects between directional and nondirectional information of auditory warning on driving behavior. Discrete Dyn. Nat. Soc. 2015, 1–7.