Complementary audio-visual collision warnings

Nicola Fricke[1] & Manfred Thüring[2]

[1] Transportation Active Safety Institute (TASI), Indianapolis, USA
[2] Technische Universität Berlin, Germany

The growing number of driver assistance systems increases the demand for warnings that are intuitively comprehensible. Particularly in hazardous situations, such as a threatening collision, a driver must understand the warning immediately. For this reason, collision warnings should convey as much information as needed to interpret the situation properly and to prepare preventive actions. The present study investigated whether informing about the object and the location of an imminent crash by a multimodal warning (visual and auditory) leads to shorter reaction times and fewer collisions compared to warning signals which only inform about the object of the crash (auditory icons) or give no additional information (simple tone). Results reveal that multimodal warnings have the potential to produce a significant advantage over unimodal signals as long as their components complement each other in a way that realistically fits the situation at hand.

INTRODUCTION

As more technology is introduced into road vehicles, new challenges arise. One of these challenges results from the variety of messages delivered by driver-assistance systems: How can we ensure that such a message (a) can be distinguished from other messages, (b) delivers the appropriate information and (c) leads to the desired effect? We will address these issues for a type of message that is of particular importance - the warning against a potential collision.

Two problems must be solved in the development of collision warnings. The first one is technical in nature. Based on data from various sensors, such as radar and laser, sophisticated algorithms are needed to automatically recognize the hazard of a collision. Research to fulfill this requirement is well under way and will not be discussed here (Kaempchen & Dietmayer, 2003). The second problem is psychological in nature and concerns the three issues outlined above. A collision warning must not be mistaken for less relevant information, must contain all crucial information needed in the situation, and must trigger adequate preventive actions on behalf of the driver.

Since collision situations are rare events, the meaning of a warning can hardly be learnt by experience. Instead, information about such hazards must be intuitively comprehensible. Moreover, the driver must respond to the warning as fast as possible. How can this be achieved?

According to Stanton and Edworthy (1999), auditory information is particularly well suited for events that require an immediate response, as long as the signal is concise. Verbal information does not comply with this criterion because its cognitive processing takes too long in time-critical situations (Campbell, Richard, Brown, & McCallum, 2007). Hence, nonverbal acoustic signals seem more appropriate as collision warnings, but unfortunately they are not as informative as verbal messages. Simple tones are not easily conceived as warnings, nor do they inform about the *nature* of the hazard or its *location*.

One way to advise the driver of the *nature* of the hazard is offered by auditory icons. They are everyday sounds most people are familiar with and therefore can understand easily and quickly (Barrass & Kramer, 1999; Petocz, Keller, & Stevens, 2008). Various studies have used them as warning signals because of their potential to reduce cognitive effort (Belz, Robinson, & Casali, 1999; Graham, 1999) and their ability to elicit fast reactions (Sanders & Mc Cormick, 1993). Applied as collision warnings, they may point out what to expect, as for instance the ringing of a bicycle bell to signal a bicycle appearing unexpectedly from behind an obstruction.

The *position* of a hazard, or the *direction* it is approaching from, can be indicated either acoustically or visually. People in a car are able to localize sounds quite accurately under static laboratory conditions (Tan & Lerner, 1996), and in studies with a driving context, spatial auditory icons led to enhanced visual target identification (Ho & Spence, 2005; Ho, Tan, & Spence, 2006). In a more realistic driving study though, participants were not able to benefit from auditory warnings conveying spatial information (Fricke & De Filippis, 2008). No advantage of these warnings in terms of reduced collision frequencies or faster brake reactions could be ascertained.

As an alternative for localizable sounds, visual cues can be employed. This may be also advantageous since other sounds inside a car can interfere with an auditory-only signal and deteriorate its spatial information. For this reason, most guidelines recommend using a multimodal warning signal (Campbell, et al., 2007), such as a combination of acoustic and visual or haptic cues.

Following this proposal, our study investigates audio-visual collision warnings. A combination of an auditory warning signal which is either specific (auditory icon) or unspecific (tone) is combined with a spatial visual LED-signal to a multimodal warning. These combinations are compared to unimodal auditory warnings.

The multimodal warnings do not merely provide redundant information – as tested in other studies, for example by coupling the sound of a horn with the visual display of a car (Belz, et al., 1999).

Figure 1: Car scenario

Figure 2: Bike scenario

Instead, the different components of information constituting the multimodal signal are complementary to each other (Milekic, 2002). The spatial information provided by the LED signal amends the information about the collision object provided by the auditory icon, both contributing to a richer and more comprehensive representation of the hazard.

Two research questions arise for such complementary warnings: (a) Is a specific warning which informs about the nature of the hazard (auditory icon) more helpful than an unspecific warning (tone)? (b) Is a multimodal signal which includes information about the location of the danger a more effective and efficient warning than a unimodal signal?

If complementary multimodal warnings were beneficial they should lead to faster reactions and thus help to reduce the number of collisions.

METHOD

The study was conducted in a fixed-base driving simulator with artificial driving noise. The simulation was projected on a screen located 1.5 meters away from the car and 1.72 meters in height as well as 2.35 meters in length.

Materials

Two different scenarios were used to simulate a potential collision. In the car scenario, a van driving ahead on a rural road suddenly braked to a complete stop (see figure 1). Since the brake lights were disabled, the participants noticed the braking very late. In the bike scenario, participants were driving alongside a row of parked cars when suddenly a bicycle entered the street (see figure 2).

To be consistent with these scenarios, two different auditory icons were used as specific warnings, i.e., screeching tires and a bicycle bell. Both auditory icons had been tested in pilot studies. Participants had been able to respond to them relatively fast (< 1100 ms) and could identify them at an average rate of more than 85% in a free association task. As an unspecific warning, a 600Hz tone (see also Graham, 1999) was employed in both scenarios. All acoustic warning signals were normalized stereo files that were played at 71-75 dB SPL through the car speakers and lasted one second (Stevens, Brennan, & Parker, 2004).

The visual warning was supposed to direct drivers' attention to the location where the collision object was about to appear. Therefore, it had to be established unobtrusively in their visual field. To accomplish this, an LED-band was positioned right under the windshield on the dashboard. A similar approach (see figure 3) had been tested successfully to indicate pedestrians' positions in night-vision enhancement systems (Mahlke, Roesler, Seifert, Krems, & Thuering, 2007).

Figure 3: visual LED warning (Mahlke et al., 2007)

The LED-band was 80 cm long and covered the driver's visual field of the screen on which the simulation was projected.

Depending on the position of the collision object, a segment of about 5 cm corresponding to the actual spatial position of the object (car or bike) was illuminated red and flashed. This lasted until the car had completely stopped. The flash-rate was 10 Hz as recommended by Sanders & McCormick (1993).

Design

Three independent variables are investigated in our study: the warning type (tone or auditory icon), the visual component (without LED, or with LED), and the sequence in which the scenarios were presented (car-bike or bike-car). The visual information was always supplementary to the auditory warning. There was no visual-only condition.

All independent variables were implemented as between-subjects factors. Accordingly, the design of the study was a 2x2x2-factorial design. The dependent variables were brake-reaction time and number of collisions.

Procedure

First, participants completed questionnaires concerning demographic factors and simulator sickness. Afterwards, they were introduced to the functions of the driving simulator and completed a training drive lasting for about five minutes. In the main part of the experiment, everybody drove two separate routes, each lasting approximately three to four minutes. On each route, they were confronted with one of the collision scenarios and got warned either by an auditory only signal (tone or auditory icon) or by a combination of auditory and visual signal (tone + LED or auditory icon + LED). All participants experienced each of the scenarios only once. The collision warning was always presented at least 3 seconds before the actual collision would occur if no preventive action were performed. At the end of the experiment, participants were interviewed to gather information on their subjective impressions of the warnings.

Participants

One hundred sixty persons participated in the experiment. All of them were men because former studies had shown that female drivers were much more prone to simulator sickness. Participants were matched for age which ranged from 18-56 (no more detailed age information can be provided due to data security). All of them had a driver's license and more than 85 % used their car every day and for more than 10.000 km per year.

RESULTS

Since the warnings in the auditory icon condition were different in the two scenarios (screeching-tires vs. bicycle bell), both scenarios were evaluated separately.

Reaction times

The factors were warning type (auditory icon vs. tone), visual component (without LED vs. with LED), and scenario sequence (bike first or bike second). Brake reaction time was defined as the interval between the appearing of the warning and braking.

In the *bike scenario* a main effect was found for the scenario sequence $F[2,121] = 22.88$, $p < .01$. When the bike scenario was presented first, brake reaction times were much longer compared to its presentation as second scenario. Another main effect resulted from the factor visual component $F[2,121] = 4.80$, $p < .05$. Multimodal warnings led to faster braking reactions than unimodal warnings.

The factor auditory warning type showed no effect, nor could any interaction effects be found. Mean values and standard errors of the main effect of the visual component are shown in figure 4.

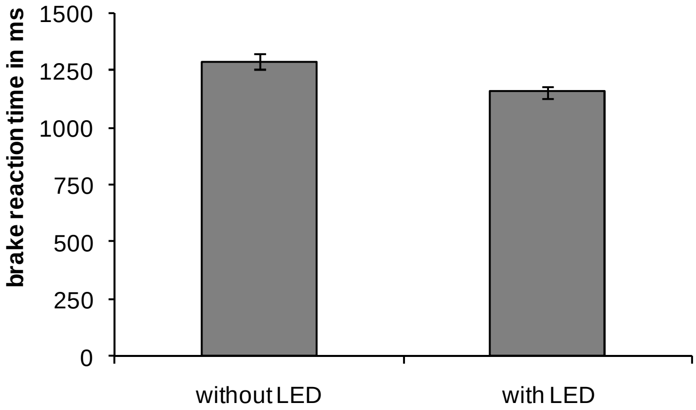

Figure 4: Mean values and standard errors of brake reaction times for the visual component effect in the bike scenario

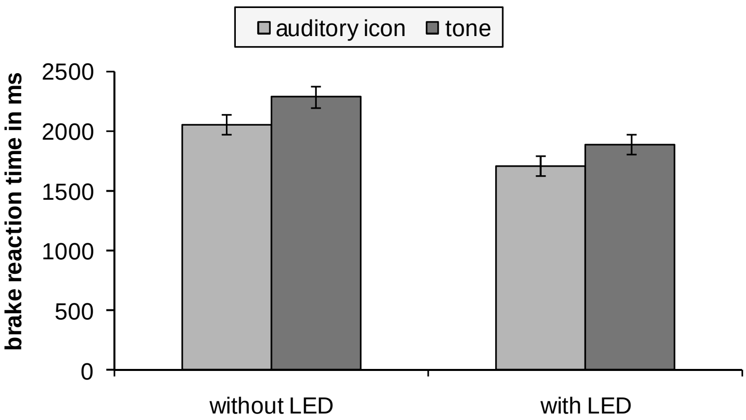

Figure 5: Mean values and standard errors of brake reaction times for the auditory warning type and visual component effect in the car scenario

In the *car scenario*, three significant effects were found. In accordance with the bike scenario, both scenario sequence $F[1,149] = 38.84$, $p < .01$, and visual component $F[1,149] = 19.03$, $p < .01$, produced a main effect. Again, reactions took longer for warnings in the first scenario and multimodal warnings reduced reaction times compared to unimodal warnings.

In addition, the auditory warning type led to a significant main effect $F[1,149] = 6.06$, $p < .05$, and the auditory icon produced faster reactions than the tone warning.

No interaction effects were found. Mean values and standard errors of the auditory warning type and modality effect are shown in figure 5.

Collision frequency

Only drivers in the car scenario caused collisions whereas all drivers in the bike scenario succeeded in avoiding any crashes. Absolute frequencies of collisions in the car scenario are displayed in table 1. A χ^2-test showed a significant main effect for the auditory warning type (Pearson $\chi^2 = 4.43$; $p < .05$) displaying a greater number of collisions for the tone compared to the auditory icon.

Table 1: Absolute frequencies of collisions for the auditory warning types in the car scenario

warning type	collision		total
	no collision	collision	
auditory icon	58	10	68
tone	50	21	71
total	108	31	139

DISCUSSION

Our results showed different effects for the two collision scenarios under investigation.

In the *bike scenario*, the multimodal warning produced faster brake reactions, no matter whether the warning included a simple tone or an auditory icon. Hence, one might conclude that the visual cue helped drivers to orient their attention. The auditory icon, on the other hand, did not lead to further improvements. For this scenario, information about the potential collision object, as provided by the bicycle bell, did not lend additional support. The reason for this lack of an effect remains unclear. It cannot be ruled out that the icon was not adequate for this scenario, i.e., that the bicycle bell did not prime the drivers in the way it had been expected. On the other hand, the results of our pre-studies make this interpretation rather unlikely. Another possibility is that the collision situation in the bike scenario was easier to handle than the one in the car scenario. Hence, the visual cue alone may have accelerated responses so much that the situation was pretty easy to cope with. Such a ceiling effect may have prevented the auditory icon from increasing drivers' performance any further. This interpretation appears more likely and is supported by the finding that no collisions at all appeared in the bike scenario.

In the *car scenario*, collisions did appear. As the analysis of collision frequencies revealed, there was a positive effect of the auditory warning type (screeching tires), but no positive influence of the visual component (red LED) could be ascertained. Two factors may have contributed to the absence of such an effect:

1. Due to individual differences in driving behavior, the distance between the drivers and the van in front of them was not a constant. As a consequence, some drivers were very close to the other car while others were further away when the warning was given. For this reason some people had more difficulties to avoid a collision than others. Obviously, this increases the error variance and reduces the chances of an effect to reach the required level of significance. The only option to avoid this problem would have been to directly instruct our participants to drive at a particular speed and to keep a predefined distance to cars in front of them. We did not choose this approach to ensure more natural driving behavior.
2. In the car scenario, the collision object was clearly visible and directly in front of the driver. Therefore, a positive effect of indicating the position of the hazard may have been much less pronounced than in the bike scenario where the collision object was still out of sight at the time of warning and then suddenly appeared at the position that the visual cue had indicated.

With respect to reaction times in the *car scenario*, we found an effect of the warning type, i.e., the auditory icon had an advantage over the tone warning. In this particular situation, information about the nature of the collision object was helpful to accelerate the drivers' responses. Moreover, there was an effect of modality. Multimodal warnings led to faster reactions than unimodal ones. This supports the assumption that visual cues play an important role in priming drivers for reacting to an impending collision. At least in the car scenario, the combined effect of information about the nature of the danger and its location led to the fastest reactions.

This finding is further supported by some participants who mentioned that the red LED light reminded them of a brake light. The combination of screeching-tires and a red LED in the car scenario probably matches the situation better than the combination of a bicycle bell and a red LED in the bike scenario. It seems that the constellation of an auditory icon and a visual cue in the car scenario constitutes a meaningful pattern which represents the hazardous situation more appropriately.

At first sight, our findings concerning reaction times in the car scenario seem to be contradictory to our findings concerning the collision frequencies where the visual cue (as part of the multimodal warning) did not help to reduce the number of collisions. This contradiction can be eliminated by assuming that the visual cue successfully contributed to decreasing reaction times, but that the resulting time savings were not big enough to significantly decrease the number of collisions.

Why was the visual cue not efficient enough to decrease collision frequency in the car scenario, and why was it of no help at all in the bike scenario? An answer to this question may result from the data of the interviews conducted at the end of the study. While people had no difficulties in noticing the LED light, only 13.8 % of the participants could tell its position and reported that the light had appeared at different locations in the two scenarios. This gives rise to the assumption that *exact* spatial information transmitted by a multimodal warning may not be as helpful as expected because people might not be able to process it in a hazardous situation that calls for a fast response. This assumption fits to the interpretation given above: Not the spatial information helped to decrease reaction times in the car scenario, but the constitution of a consistent pattern that represents the hazardous situation in a meaningful way. Of course, this interpretation must be verified in future studies.

A third effect, we found in our study, concerned the sequence in which participants were exposed to the two different collision situations. For both scenario types, the second confrontation with a collision warning led to faster responses. This indicates a pronounced learning effect. As soon as participants had experienced the first collision scenario, they expected similar events for the continuation of the experiment and were better prepared to react. This finding has methodological implications that are particularly important for future studies with respect to ecological validity. In real-life situations collision warnings would be rare events and drivers would not expect them. Empirical studies should take this into account and restrict the number of warning trials drastically. It might even be most appropriate to confine such studies to one warning trial because any data produced in subsequent trials might be heavily biased by learning and expectations.

To summarize, let us answer the two research questions addressed in this study. Auditory icons which inform about the nature of a collision hazard can be more helpful than unspecific warnings, such as a simple tone. Moreover, multimodal warnings which provide visual cues indicating the location of the collision object can be more helpful than unimodal warnings consisting of mere acoustic information. Please note our claim that both measures "*can be* helpful". If they really *are* helpful depends on whether they are combined into complementary multimodal warnings which realistically accord with the situation at hand - as in the case of coupling the sound of screeching tires with the red light of an LED to warn about the sudden braking of a car in front.

ACKNOWLEDGEMENTS

This research was performed while the first author was on appointment at Technische Universität Berlin.

REFERENCES

Barrass, S., & Kramer, G. (1999). Using sonification. *Multimedia Systems*, *7*(1), 23-31.

Belz, S. M., Robinson, G. S., & Casali, J. G. (1999). A new class of auditory warning signals for complex systems: Auditory icons. *Human Factors, 41*(4), 608-618.

Campbell, J. L., Richard, C. M., Brown, J. L., & McCallum, M. (2007). *Crash warning system interfaces: Human factors insights and lessons learned (Report No. DOT HS 810 697)*. Seattle: Battelle Center for Human Performance and Safety.

Fricke, N., & De Filippis, M. (2008). Effects of auditory warnings on driving behaviour. In D. d. Waard, F. O. Flemisch, B. Lorenz, H. Oberheid & K. A. Brookhuis (Eds.), *Human Factors for assistance and automation* (pp. 117 - 128). Maastricht: Shaker.

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

Ho, C., & Spence, C. (2005). Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *Journal of Experimental Psychology: Applied*, *11*, 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.

Kaempchen, N., & Dietmayer, K. (2003). *Data synchronization strategies for multi-sensor fusion.* Paper presented at the 10th World Congress on Intelligent Transport Systems (ITS 2003), Madrid, Spain.

Mahlke, S., Roesler, D., Seifert, K., Krems, J. F., & Thuering, M. (2007). Evaluation of six night vision enhancement systems: Qualitative and quantitative support for intelligent image processing. *Human Factors*, *49*(3), 518-531.

Milekic, S. (2002). *Towards tangible virtualities: Tangialities*. Retrieved November 17, 2008, from http://www.archimuse.com/mw2002/papers/milekic/milekic.html.

Petocz, A., Keller, P. E., & Stevens, C. (2008). Auditory warnings, signal-referent relations and natural indicators: Re-thinking theory and application. *Journal of Experimental Psychology: Applied*, *14*, 165-178.

Sanders, M. S., & Mc Cormick, E. J. (1993). *Human factors in engineering and design*. New York: Mc Graw-Hill.

Stanton, N. A., & Edworthy, J. (1999). Auditory warning affordances. In N. A. Stanton & J. Edworthy (Eds.), *Human factors in auditory warnings* (pp. 113-128). Aldershot: Ashgate.

Stevens, C., Brennan, D., & Parker, S. (2004). Simultaneous manipulation of parameters of auditory icons to convey direction, size, and distance: Effects on recognition and interpretation. *Proceedings of the tenth Meeting of the International Conference on Auditory Display*.

Tan, A. K., & Lerner, N. D. (1996). *Acoustic localization of in-vehicle crash avoidance as a cue to hazard direction*. Washington, DC: National Highway Traffic Safety Administration.