Comparing an intelligent and a conventional headway-based auditory feedback system on safety and acceptance in an on-road car following experiment

Coert de Koning, Hidde Lingmont, Tjebbe de Lint, Florian van den Ouden, Twan van der Sijs

***Abstract*— Auditory feedback produced by advanced driver assistance systems is often perceived as annoying, which may result in misuse of these systems [1] [2]. An intelligent time headway-based auditory feedback system is designed to improve user's acceptance and to improve driving safety. The intelligent feedback system provides feedback depending on the 'danger zone' and the duration of remaining in said danger zone. Additionally, the intelligent system does not provide feedback when time headway is increasing in order to reduce the number of false alarms. An on-road experiment is conducted to investigate the acceptance and driving safety of the intelligent system compared to a Mobileye-based system. A Mobileye-based system is regarded as conventional and uses only one feedback threshold. 20 participants drove two trials on the highway, one with a system that resembles the Mobileye system and one with an intelligent system. After the experiment, the participants completed a questionnaire to assess their opinion about the usefulness and satisfaction of both systems. Sensory data were collected from the vehicle and used to evaluate the safety of both systems. Safety is defined as the percentage that the time headway is less than 0.6 seconds relative to the time the time headway is less than 2 seconds. The experiment shows that the trials in which the intelligent system was used scored better on safety than the Mobileye-based system. However, the difference in safety between both systems is not statistically significant. The intelligent system was also perceived as more satisfying, where the Mobileye-based system was found to be more useful. The difference between both systems on usefulness and satisfaction were not statistically significant. The intelligent system has the potential of being safer without compromising acceptance, hence more research on this system could be useful.**

***Index Terms*— Acceptance, Annoyance, Auditory feedback, Following Distance, On-Road Experiment, Safety, Time Headway, Vehicle Warning Systems**

I. INTRODUCTION

DUE to the rise of feedback systems in cars, driving is safer than ever before [3]. In today's cars, there are multiple ways to give feedback to the driver. Some systems provide either visual, haptic or auditory feedback whereas other systems use a combination of these three.

On the aspect of auditory feedback, a substantial amount of research has already been done. Two possible ways to give auditory feedback are non-speech and speech-based. A benefit of auditory feedback is that it has the possibility to transfer messages that convey various urgency levels. For example, two-tone chimes appeared to be a good sound for cautionary warnings whereas speech-based feedback was the best for transferring complex messages according to Campbell et al. [4]. A benefit of speech-based feedback is the possibility to communicate semantically rich information to the driver. It has been found that a woman's voice is preferred over a man's voice when receiving feedback [5]. A potential disadvantage of using auditory feedback is that it may annoy drivers. Research on this topic has concluded that an annoyance trade-off should be made when designing a warning feedback system [6] [7].

Options for improving safe driving are user-initiated individualisation and system-based individualisation, which are both forms of personalised feedback. A system-based personalised feedback system was found to have a favourable outcome on acceptance compared to a non-personalised feedback system [8]. Additionally, an Forward Collision Warning system adapted on the user's reaction time was found to have benefits on acceptance when the driver has an aggressive driving style but provided no significant advantages for non-aggressive drivers compared to a non-adaptive feedback system [9].

Intelligent feedback is not standard in present-day cars, and a multiple state auditory feedback system could improve safe driving since different urgency levels require different types of feedback [4]. Therefore, the following research question is addressed in this paper: 'Is a real-time auditory feedback system on time headway safer and better accepted if it adjusts its feedback on headway thresholds and the driver's response?'. It has been tested whether such an intelligent real-time time headway-based auditory feedback system improves safety in traffic and increases acceptance, compared to a conventional system. If at least the same acceptance can be maintained, the system has a chance of getting accepted in a realistic driving environment.

The hypothesis is that the use of intelligent feedback concerning headway and the driver's response to real-time auditory feedback can improve safe driving in passenger cars compared to the use of the Mobileye-based system. This is based on the expectation that the driver is more likely to accept warnings earlier if the warnings sound more friendly. Therefore, the first feedback can be given at a greater time headway, which might improve the safety. The drivers might then become more aware of headway danger and adapt their behaviour based on the feedback provided.

To test whether the hypothesis is correct, two experiments were conducted. In the first experiment, it was tested which sound is most suitable for giving the first level of feedback and at which headway time this feedback should be provided. The results are used in the design of the multiple state feedback system, which was tested in the second experiment. From the results of this second experiment, a conclusion could be drawn on whether the hypothesis is true.

II. Methodology

Prior to the main experiment, a lab experiment was conducted to determine the feedback signal for the first headway warning and the preferred time headway (THW) at which the first headway warning is given. The lab experiment was an online survey with 67 participants. The resulting preferred time headway for the first feedback was the shortest of the three options, namely 0.5 seconds. Taking literature on safe following distances into consideration [10], it was decided to give the initial feedback at a time headway of 0.8 seconds. A more extensive description of the lab experiment can be found in the supplementary information, together with a more elaborate explanation of the analysis of the results.

The main experiment is constructed to measure the safety and acceptance of an intelligent system and a conventional system. This is done by letting a group of 20 volunteers take a drive in a test car with both systems. Only the participants that received feedback from both systems will be considered in the analysis. Of the 11 participants that received feedback, two were female, and 9 were male. All participants were between 18 and 26 years old. 1 out of the 11 useful participants drives less than 1000 km per year, 8 drive 1001 to 5000 km per year and 2 participants drive 5001 to 15000 km per year.

A. Hardware

A Volvo C30 equipped with a Mobileye camera was used as the test car and a Nissan Micra was used as the leading car. The auditory feedback was given using a speaker. The raw data were processed on a Raspberry Pi with software, converting the raw Mobileye data to numeric values, provided by SD-Insights [11].

B. The systems

a) The conventional system: The conventional system (System C) gives a feedback sound if the driver is in a situation where the time headway is below 0.6 seconds.

b) The intelligent system: The intelligent system (System I) uses different states for three situations.

1) 0.5 s $<$ THW ≤ 0.8 s: The sound resulting from the lab experiment (Sound 1) will be played the first time the driver arrives under the threshold of 0.5 seconds. After that first feedback instance, if the driver remains in the time headway interval between 0.8 s and 0.5 s, every 8 seconds a friendly female voice saying in Dutch 'volgafstand te kort' (Voice 1) will be played. This phrase translates to: 'following distance too short'.
2) 0.3 s $<$ THW ≤ 0.5 s: A more urgent variant of Sound 1 (Sound 2) will be played the first time the driver arrives under the threshold of 0.5 seconds. If the driver remains in the time headway interval between 0.5 s and 0.3 s for more than 5 seconds, a less-friendly female voice saying in Dutch 'neem meer afstand' (Voice 2) will be played every 5 seconds until the time headway becomes larger than 0.5 seconds. This phrase translates to: 'increase distance'.
3) THW ≤ 0.3 s: The most urgent sound (Sound 3) will be played continuously until the driver increases his time headway above 0.3 seconds.

System I also calculates a 'reaction'-value, which aims to recognise the drivers' response to their situation and tries to reduce the false positives that occur for merging traffic. The system only gives feedback if the drivers do not react to their situation. The reaction value uses the relative speed between the vehicle and the vehicle in front. It also uses the first difference, with a sampling frequency of approximately 10 Hz, of the time headway itself. If both these values are positive, Sound 1 and Sound 2 will be suppressed, and the voices and Sound 3 will not be suppressed. Further explanation of the feedback algorithm and the sound files can be found in the supplementary information.

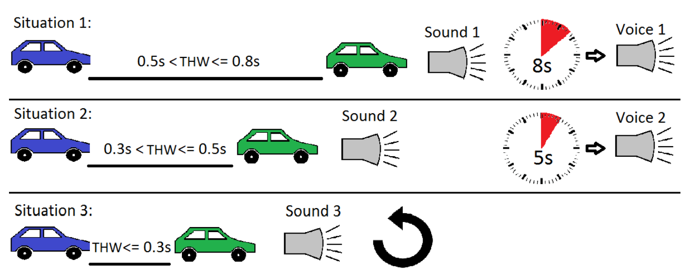

Fig. 1: A visualisation of System I. A sound is played once when the car comes within a time headway of 0.8 s. When the car stays in the state between 0.5 s - 0.8 s THW, Voice 1 is played. When the car crosses the 0.5 s THW threshold, Sound 2 is played once. When the car stays in the state between 0.5 s - 0.8 s THW, Voice 2 is played. When the car crosses the 0.3 s THW threshold, Sound 3 is played continuously until the car is out of this zone.

C. Procedure

After signing a form of consent, the participants first completed a questionnaire consisting of a DBQ and questions to determine their driving background. The participants then took place in the test car, provided by SD-Insights, together with two observers. A given route was driven, of which a map is provided in the supplementary information. The driver was not informed about the route but was instructed to follow a leading car driven by one of the authors. Half of the experiments started with System I enabled and the other half started with System C enabled. Halfway, the feedback system was changed from System I to System C or vice versa. The driver was notified of the system change. When the car had returned to its starting location, the driver was asked to complete a questionnaire to measure acceptance [12].

Afterwards, an interview covering design parameters (thresholds, sounds and voices) to validate the design of the system was conducted. In this interview, the participants were asked to compare the two systems and tell which one they preferred. It was also asked what they thought of the set-up of System I and whether they liked this system. The data from the drive were collected from the car to determine safety. The form of consent, the DBQ and the acceptance-questionnaire can be found in the supplementary information.

It was decided to let the participant drive themselves, because people may react differently to their own mistakes than to other peoples mistakes. This could alter their acceptance of the feedback. Driving in reality, compared to a simulation, is preferred because participants are known to drive somewhat differently in road situations as compared to a simulator [13].

D. Foreseen pitfalls

A possible problem could be that the participants want to leave a good impression and therefore maintain a time headway that is bigger than usual. This could pose a problem when the driver spends too little or no time in an increased risk zone to get feedback. When drivers do not get any feedback, their drives are not useful for this research. For this reason, the driver needs to follow a leading car, since in general, this lowers the average time headway [14]. Following a car has the effect that drivers spend more time at an unsafe distance [3] [15] [16] [14]. This means more useful data can be obtained in a shorter amount of time.

A disadvantage of an on-road experiment is the uncontrolled environment in which the experiment is conducted. An example of this is that there could be a traffic jam. This could cause irritation for the driver along with other changed parameters. To minimise the chance of getting in a traffic jam, alternative routes were made and used 7 times in the experiment.

E. Analysis

To express danger, a danger factor is computed for each participant for both systems. The danger factor is the percentage of the time the driver spends with less than 0.6 seconds time headway out of the time the driver spends under 2.0 seconds time headway (eq. 1). A lower danger factor represents a participant driving safer.

$$\text{Danger Factor} = \frac{\text{THW} < 0.6\text{ s}}{\text{THW} < 2.0\text{ s}} \cdot 100\% \qquad (1)$$

The danger factor indicates how much danger the driver has been in during their driving time. The averages for the intelligent and the conventional feedback systems can be compared, which could be used to draw conclusions about safety. The acceptance questionnaire gives insight in acceptance of the two systems by the drivers.

III. Results

During the experiments, 11 out of the 20 participants got feedback from both systems. Only the results of these 11 participants will be considered here.

A. Time headway data analysis

The gathered time headway data for both systems is plotted in a normalised histogram in Figures 2 and 3. As can be seen, the mean time headway for System C is 1.47 s, whereas the mean time headway for System I is 1.55 s.

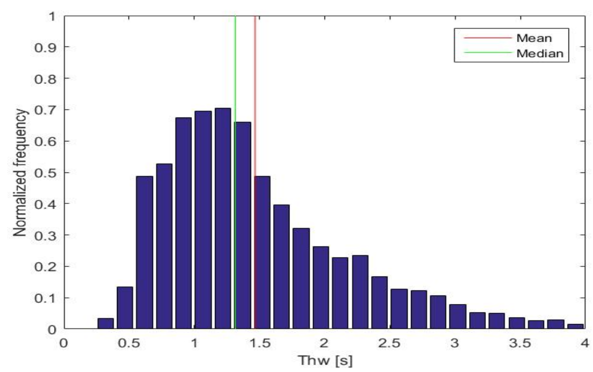

Fig. 2: A normalised histogram of the THW for System C with mean (THW = 1.47 s) and median (THW = 1.31 s).

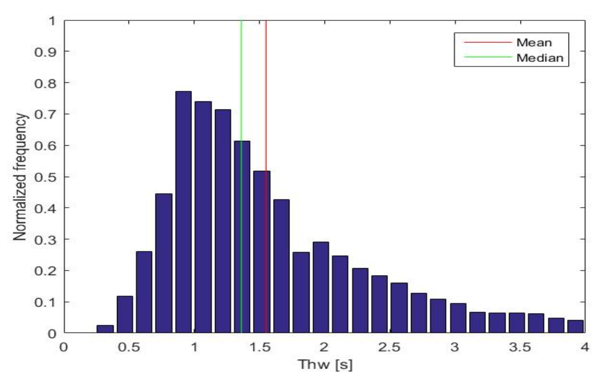

Fig. 3: A normalised histogram of the THW for System I with mean (THW = 1.55 s) and median (THW = 1.36 s).

For System C, the percentage of time that the time headway was less than 2.0 seconds relative to the total driving time, was 53%. For System I, this value is 49%. The time headway that was greater than 2 seconds is not used for the analysis. The assumption is that the vehicle is following a vehicle if the time headway is less than 2.0 seconds.

In Table I, the distribution of different percentages of time the time headway was below 2.0 seconds is shown in three bins together with the danger factor. Between 2.0 seconds and

1.2 seconds, between 1.2 seconds and 0.8 seconds and a time headway less than 0.8 seconds.

TABLE I: Distribution of all THW values < 2 s for System C and System I

THW [s]	System C [%]	System I [%]
> 1.2	46.8	48.2
$1.2 - 0.8$	34.2	38.1
< 0.8	19.0	13.6
< 0.6 (Danger Factor)	5.3	4.1

Furthermore, plots are made of all the instances the danger threshold of 0.8 seconds and 0.6 seconds were crossed, together with an average time headway development after crossing this threshold for both systems. These plots can be found in Figures 4, 5, 6 and 7. For these plots, it was assumed that the time between the data points is 0.1 seconds and equal for each step. There were 30 instances for which the 0.8 s threshold was crossed for System C and 27 for System I. The 0.6 s threshold was crossed with system I at 5 instances and with system C at 28 instances. All the lane changing manoeuvres are filtered because only approaches were considered.

The development has been plotted from 1 second before to 5 seconds after crossing the threshold for both the 0.8 seconds and the 0.6 seconds approach. The difference between the mean values of the two systems, 5 seconds after the threshold is crossed, is not statistically significant at the 0.8 seconds approach. A t-test for unpaired samples on these values gives that System I scored higher (M=0.88, SD=0.19) than System C (M=0.79, SD=0.14), $t(55)=-1.89$, p=0.064, 95% CI [-0.17, 0.01]. The 0.6 seconds approach does also not show a statistically significant result. A t-test for unpaired samples on these values gives that System C scored higher (M=0.82, SD=0.26) than System I (M=0.62, SD=0.094), $t(31)=1.62$, p=0.116, 95% CI [-0.05, 0.44]

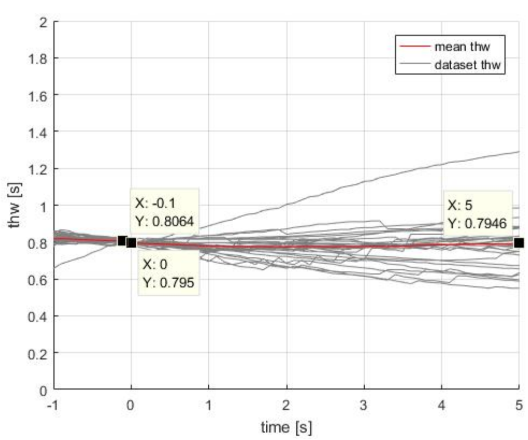

Fig. 4: In this figure, time headway developments for all approaches with System C are plotted. At time = 0 s, the 0.8 s threshold is crossed.

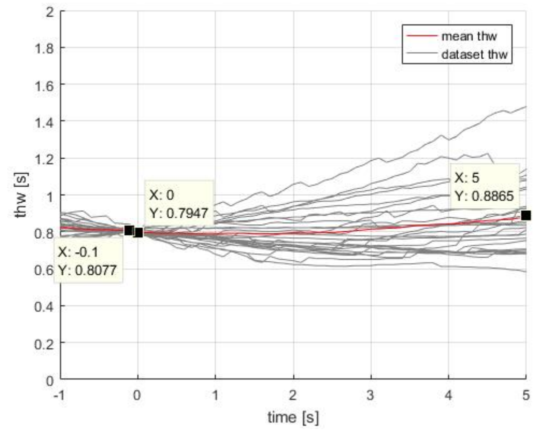

Fig. 5: In this figure, time headway developments for all approaches with System I are plotted. At time = 0 s, the 0.8 s threshold is crossed.

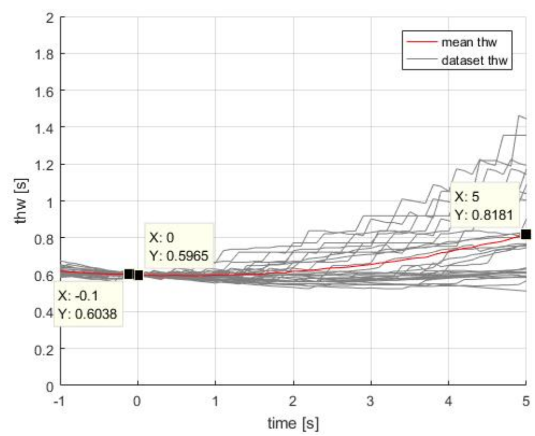

Fig. 6: In this figure, time headway developments for all approaches with System C are plotted. At time = 0 s, the 0.6 s threshold is crossed.

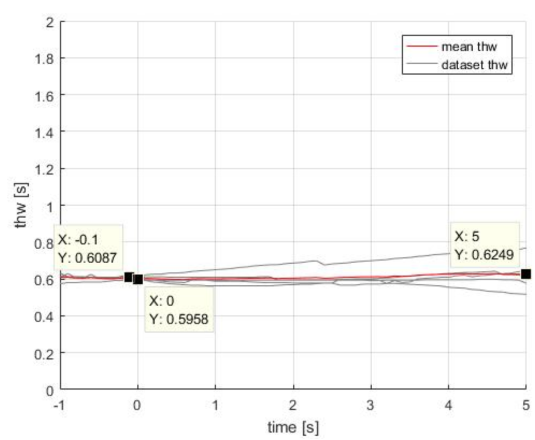

Fig. 7: In this figure, time headway developments for all approaches with System I are plotted. At time = 0 s, the 0.6 s threshold is crossed.

B. Danger factor calculation

In Figure 8, box plots are shown for the two systems. System I scores better on the danger factor, with some outliers, while the danger was generally higher with less outliers with System C. A statistical t-test with $H_0 : \mu_I = \mu_C$ and $H_1 : \mu_I \neq \mu_C$ over the danger factors gives that System C scored higher (*M*=5.8, *SD*=4.9) than System I (*M*=4.3, *SD*=8.0), *t*(10)=0.706, *p*=0.496, 95% CI [-3.09, 5.95].

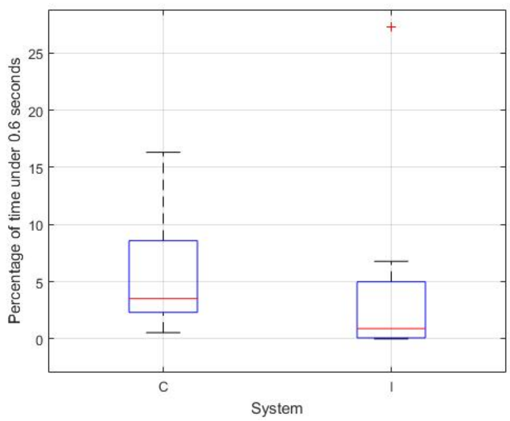

Fig. 8: A box plot of the danger factors for both systems.

C. Acceptance

1) Questionnaire: The acceptance questionnaire measures two parameters, namely the satisfaction and the usefulness of the systems. The average results together with the standard deviations and Cronbach's alphas are presented in Table II. By doing a t-test for paired samples, the difference in satisfaction and usefulness between the two systems were calculated. For the usefulness System C was found larger than System I with *t*(10)=0.978, *p*=0.351, 95% CI [-0.33, 0.83]. For the difference in average satisfaction System I scored higher than System C with *t*(10)=-1.177, *p*=0.266, 95% CI [-1.18, 0.37].

2) Interview: In the interview, contradictory opinions came forward. For example, the participants disagreed on the helpfulness of the repeating voices and Sound 1. There were only two items that came back more frequently for System I. Voice 1 was found to be hard to understand by 4 people and the first threshold for the provided feedback was found to be too early. The majority of participants would like to receive feedback at a shorter time headway. This corresponds with the results of the lab experiment.

TABLE II: The average usefulness and satisfaction for both System C and System I on a scale from -2 to 2 based on the acceptance questionnaire.

	Mean usefulness	σ usefulness	Reliability coefficient usefulness	Mean satisfaction	σ satisfaction	Reliability coefficient satisfaction
System C	0.22	0.90	0.83	-0.34	1.01	0.89
System I	-0.04	0.98	0.46	0.07	0.97	0.89

IV. Discussion

This study was conducted to find out whether the intelligent system performs better on safety and acceptance than the conventional system. To verify this, an on-road experiment was conducted, as explained in the methodology section.

Two histograms were made to analyse the results, as can be seen in Figures 2 and 3. The mean and median of the time headway differ little for System C and System I. This implies that there is no big difference between the two systems. The histogram of System C, however, shows a drop round 0.6 seconds time headway, where the histogram of System I shows a drop around 0.8 seconds. This might be the result of the different time headway at which the two systems provide the first feedback.

In Table I, the percentage the driver spends in a certain time headway interval is given, when only taking into account the amount of time the driver had a time headway less than 2 seconds. An observation that can be made from this table is that participants spend a lower percentage of time with a time headway smaller than 0.8 seconds when driving with System I. This could indicate that System I is safer than System C, because the participants spend less time in the dangerous zone of a time headway under the 0.8 seconds. The fact that System I already gives feedback at 0.8 seconds and System C only gives feedback at 0.6 seconds could be the cause of this difference.

Also, when comparing the two plots in Figures 4 and 5, it can be seen that the mean after 5 seconds is higher for System I, than for System C. Although these results are not statistically significant, it suggests that System I results in a larger time headway after feedback is given than System C. The reason for this difference could be that feedback from System I is given at 0.8 seconds, while the feedback from System C is given at 0.6 seconds.

This could also explain that when comparing the two plots in Figures 6 and 7, the mean of System C is higher after 5 seconds in comparison with the mean of System I. The results are however not statistically significant. This suggests that the participants react to the feedback given by System C by increasing their time headway, while the time headway of System I does not increase after reaching the threshold of 0.6 seconds.

When considering the danger factors, shown in Figure 8, a paired sample t-test was conducted to check the difference between the danger factors. Since the p-value is large, it cannot be concluded that one of the systems works better than the other with respect to safety.

On acceptance, in terms of perceived usefulness and satisfaction as displayed in Table II, a t-test for paired samples was used. The results showed that it could not be concluded that there was a difference between the two systems.

A. *Potential pitfalls*

Since the sample consisted mostly of students, the results might not be representative of the whole population. However, the beneficial effects of time headway feedback seem to be quite stable across gender and age [3]. So, the participant group might still give representative results for all gender and ages categories.

Another reason why the result might not be representative of the whole population is the fact that only a small participant group received feedback from both systems. There was one outlier with respect to the danger factor as can be seen in Figure 8. For both systems, the same participant was responsible for very high values in the danger factors, although no unexpected events occurred during the experiment. Since only 11 participants were taken into account during the analysis, the results were heavily influenced by this exceptional individual.

Another potential problem could be the length of the drive in the experiment. All the participants drove approximately 8 minutes on the way to the turning point and also approximately 8 minutes on the way back. If they made a longer drive with both systems, their perceptions of the systems might have been different because of habituation and possible irritation.

Also, the busyness on the road could have influenced the participants' opinions on the systems. This is not just about the difference in busyness for different participants, but also about the difference in busyness on the way back and forth for one participant. The percentage of time the participant drove with a time headway less than 2.0 seconds relative to the total drive time, differs only by 4 percentage points between the two systems. This might suggest that the effect of busyness on the road on the way back and forth did not have a large influence. Although this effect is expected to be small, the difference in busyness on the road between participants could still have influenced the opinion of the participants on the systems.

Another factor regarding the drive is that not everybody has driven an equal distance with both systems. Since sometimes there was heavy traffic, alternative routes were driven in these time-slots.

Lastly, there is a possibility that the participants drove differently, because they possibly wanted to please the researchers, felt watched during driving, or actively looked for the sound signals. The fact that they drove in a different car for the first time could also have caused different driving behaviour.

B. *Future research*

Although the p-values generated by the conducted t-test on the results are not statistically significant, the results do imply potential improvements for real-time auditory feedback systems. For future research on this topic, we advise using a large group of participants to minimise the probability that the sample of the population is not representative. It might also be better to make longer drives with the systems so the participants can shape their opinion on the long run. One way to achieve this, is by implementing the system in their own cars, so they can test how the systems work without taking any of their time. This would also immediately eliminate the pitfall that people might drive differently in a new car to them. This way of testing also minimises the effect that people might drive differently when investigators are in the car as well. Another benefit of having the system built in in people's own cars, would be that traffic jams will have less impact on their opinions. After the experiment was conducted, an improvement on the intelligent system has been made which can be found in the supplementary information. Using this version in a future experiment would be interesting.

V. Conclusion

All in all, the results from this research are slightly in favour of System I, although this cannot be proven to be statistically significant by t-tests. The gathered data on acceptance does not show one system to be the most useful or satisfying. From the approach figures and histograms, it could, however, be seen that people react to melodious feedback at either 0.8 seconds and 0.6 seconds time headway. Providing feedback at 0.8 seconds compared to 0.6 seconds means that the driver is warned earlier and therefore it has the potential of being a safer system. Since no significant difference in the acceptance was found, this could mean that it could be done without compromising the acceptance. More research on this system should be done to draw a definite conclusion.

Acknowledgement

We would like to thank SD-Insights, for providing us with their test vehicle, and for their support and knowledge during this bachelor project. Furthermore, we would like to thank Jork Stapel, Joost de Winter and Pavlo Bazilinskyy for supervising the project and providing us with their advice. At last, we would like to thank Karolien Moll, for providing us with the second, leading, vehicle.

Bibliography

[1] C. Gonzalez, B. A. Lewis, D. M. Roberts, S. M. Pratt, and C. L. Baldwin, "Perceived urgency and annoyance of auditory alerts in a driving context," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 56, no. 1, pp. 1684–1687, 2012. [Online]. Available: http://dx.doi.org/10.1177/1071181312561337

[2] R. Parasuraman and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 39, no. 2, pp. 230–253, 1997.

[3] D. Shinar and E. Schechtman, "Headway feedback improves intervehicular distance: A field study," *Human Factors*, vol. 44, no. 3, pp. 474–481, 2002.

[4] J. L. Campbell, C. M. Richard, J. L. Brown, and M. McCallum, "Crash warning system interfaces: human factors insights and lessons learned," *Final Report DOT HS 810*, vol. 697, 2007.

[5] P. Bazilinskyy and J. de Winter, "Auditory interfaces in automated driving: an international survey," *PeerJ Computer Science*, vol. 1, p. 13, 2015, peerJ Inc.

[6] E. E. Wiese and J. D. Lee, "Auditory alerts for in-vehicle information systems: The effects of temporal conflict and sound parameters on driver attitudes and performance," *Ergonomics*, vol. 47, no. 9, pp. 965–986, 2004, taylor & Francis.

[7] R. Parasuraman, P. A. Hancock, and O. Olofinboba, "Alarm effectiveness in driver-centred collision-warning systems," *Ergonomics*, vol. 40, no. 3, pp. 390–399, 1997, pMID: 9118938. [Online]. Available: http://dx.doi.org/10.1080/001401397188224

[8] J. Wang, L. Zhang, D. Zhang, and K. Li, "An adaptive longitudinal driving assistance system based on driver characteristics," *IEEE Transactions On Intelligent Transportation Systems*, vol. 14, no. 1, 2013.

[9] A. H. Jamson, F. C. Lai, and O. M. Carsten, "Potential benefits of an adaptive forward collision warning system," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 4, pp. 471–484, 2008.

[10] M. Green, ""how long does it take to stop?" methodological analysis of driver perception-brake times," *TRANSPORTATION HUMAN FACTORS*, vol. 2, no. 3, pp. 195–216, 2000, lawrence Erlbaum Associates.

[11] Sd-insights. [Online]. Available: http://www.sd-insights.eu/

[12] J. D. Van Der Laan, A. Heino, and D. De Waard, "A simple procedure for the assessment of acceptance of advanced transport telematics," *Transportation Research Part C: Emerging Technologies*, vol. 5, no. 1, pp. 1–10, 1997.

[13] A. Bittner, J. Ozgur Simsek, W. Levison, and J. Campbell, "On-road versus simulator data in driver model development driver performance model experience," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1803, pp. 38–44, 2002. [Online]. Available: http://dx.doi.org/10.3141/1803-06

[14] M. Taieb-Maimon and D. Shinar, "Minimum and comfortable driving headways: Reality versus perception," *Human Factors*, vol. 43, no. 1, pp. 159–172, 2001, pMID: 11474761. [Online]. Available: http://dx.doi.org/10.1518/001872001775992543

[15] S.-K. Chen, "Estimation of car-following safety : application to the design of intelligent cruise control," Ph.D. dissertation, Massachusetts Institute of Technology. Dept. of Mechanical Engineering, 1996.

[16] L. Evans, *Traffic Safety and the Driver*. Van Nostrand Reinhold, 1991. [Online]. Available: https://books.google.nl/books?id=OyuGJJ6rKQ0C