

***FIELD TEST RESULTS OF A ROAD DEPARTURE CRASH WARNING SYSTEM:  
DRIVER ACCEPTANCE, PERCEIVED UTILITY AND WILLINGNESS TO PURCHASE***

James R. Sayer, David J. LeBlanc, Mary Lynn Mefford, Joel Devonshire  
University of Michigan Transportation Research Institute  
Ann Arbor, Michigan, USA  
E-mail: jimsayer@umich.edu

**Summary:** The road departure crash warning system field operational test (RDCW FOT) was conducted to assess the safety impacts, driver acceptance levels, and the maturity of road departure crash warning systems as installed on a light vehicle platform. This paper presents the results on driver acceptance and perceived utility of a road departure crash warning system.

## **INTRODUCTION**

Single-vehicle road departures account for approximately one in five police-reported crashes in the U.S., but are responsible for approximately 40% of the highway fatalities (Emery et al., 2005). Two types of road departure events account for about half of the police-reported crashes: drivers inadvertently allowing their vehicle to drift off the road edge, and drivers entering curves too quickly. These events may be reduced or mitigated using an integration of two crash warning technologies that are available or in development in the passenger car market: lateral drift warning (LDW) and curve-speed warning (CSW). A field operational test was conducted to assess the potential safety impacts, driver acceptance levels, and system maturity of an integrated LDW and CSW set. This test was conducted by the University of Michigan Transportation Institute (UMTRI) within its partnership with Visteon Corporation and AssistWare Technologies (now part of Cognex Corporation), within a cooperative agreement with the U.S. Department of Transportation.

This paper presents the results of driver acceptance, perceived utility, and willingness to purchase a road departure crash warning system. A companion paper at this conference addresses the system description, experimental design, performance of the road departure system in naturalistic use, and analyses of safety impacts of the technology (LeBlanc et al., 2007). A comprehensive report on these and other topics is also available (LeBlanc et al., 2006).

## **METHODOLOGY**

The road departure crash warning system field operational test (RDCW FOT) was conducted using a fleet of 11 passenger vehicles equipped with technologies that provided LDW and CSW functions. The technology was developed and integrated onboard a vehicle fleet by Visteon and AssistWare/Cognex. Each vehicle was equipped with a data acquisition system to capture the driver's experiences using hundreds of data signals, two video streams, and audio. The vehicle platform was a 2003 Nissan Altima 3.5 SE vehicle. (Although Nissan provided engineering information to facilitate system integration, Nissan was not involved in the project and there is no relationship between the tested systems and Nissan's systems.)

A set of drivers was recruited from the general driving population of southeastern Michigan, with gender split evenly and age cells of equal size containing drivers in their 20s, 40s, and 60s. Each test participant was trained on the RDCW system, asked to complete several pre-drive subjective instruments, given a demonstration drive, and then allowed to use the test vehicle as their own for four weeks. The first week was baseline driving where driver alerts were not presented, but onboard data were captured and collected. In the three remaining weeks, the RDCW system issued driver alerts. Upon completion of the four weeks of driving, the participants returned the vehicles and completed a set of subjective instruments including questionnaires, debriefing interviews, and a joint review with a researcher using video captured during their driving experience. Some drivers returned later to participate in one of four two-hour focus groups. The 78 drivers participating in the study had a travel distance exceeding 83,000 miles.

Drivers' subjective assessments of the RDCW system were obtained through several different mechanisms. These included the post-drive questionnaire (representing the bulk of the subjective data), utility ratings of specific alerts (elicited during the debriefing sessions), and a series of focus groups. This data provides a sense of the extent to which drivers accepted the RDCW system and what changes they might want to see in future implementations. All 78 drivers completed the post-drive questionnaire and reviewed some of their warning events upon returning the vehicle. However, specific analyses may not include all 78 drivers (due to questionnaire items that were skipped, the fact that some drivers did not receive every type of warning, etc.). In addition, the focus groups involved a subset of 25 of the 78 drivers, although the strength of focus group data come not so much from the number of participants, but rather from the depth and breadth of the subjective response.

## **QUESTIONNAIRE RESULTS**

Due to the number of questions contained in the final questionnaire, only a subset of the data is presented here. Although there is currently no standardized measure of driver acceptance for new automotive technologies, one particular scale offers a step in this direction. This measure, first described in Van der Laan, Heino, and De Waard (1997), uses a five-point scale to assess nine attributes of a technology. Because the scale is sufficiently broad, researchers can use it to directly compare the acceptance of different technologies across studies. In the RDCW study, the Van der Laan scale was used to measure acceptance of both LDW and CSW, as well as of the RDCW system as a whole. The scale was integrated into the post-drive questionnaire near the end of each subsystem section (i.e., the scale appeared three times within the questionnaire). Results from the present study are presented, as are comparisons with results from another field operational test of forward collision warning and adaptive cruise control.

Each item on the Van der Laan scale is anchored by two polar adjectives, such as good and bad, and drivers are asked to rate their perception of the technology by marking a box along a continuum between the two poles. Most of the adjective pairs are presented such that the positive adjective is on the left, although a few items present the positive adjective on the right. The scale is usually scored from -2 to +2, with higher numbers corresponding to values closer to the positive adjectives and vice versa. For example, a mark in the leftmost box would be scored as +2. The nine adjective pairs are: useful—useless, pleasant—unpleasant, good—bad, nice—

annoying, effective—superfluous, likeable—irritating, assisting—worthless, desirable—undesirable, and raising alertness—sleep inducing. References to scale item numbers (see paragraph below) refer to these nine adjective pairs, in the order that they are presented above.

A series of principal component analyses carried out by Van der Laan et al. suggests that the scale can usually be reliably reduced to two components: a usefulness composite measure (consisting of items 1, 3, 5, 7, and 9) and a satisfaction composite measure (consisting of items 2, 4, 6, and 8). The authors also provide some guidelines for how to use the scale and analyze the results, such as how to assess whether the two components fit a particular set of data. They first suggest using scale reliability analyses (e.g., Cronbach's alpha) to determine how well items in each component correlate with each other. Their recommended criterion for the Cronbach's alpha is at least 0.65 for each component. They suggest averaging the component scores for each subject to arrive at one usefulness and one satisfaction score for each subject. These two scores, averaged across subjects, represent the overall perceptions of usefulness and satisfaction associated with the technology. Positive numbers correspond to positive perceptions, and negative numbers correspond to negative perceptions. All of these recommended steps were carried out for the three Van der Laan scales in the present study (RDCW, LDW, and CSW).

*RDCW.* Scale reliability tests were run for the RDCW components of usefulness and satisfaction. Cronbach's alphas for RDCW usefulness and satisfaction were 0.88 and 0.86, respectively. The usefulness component had a mean score of 1.23 (SD = 0.76), which indicates positive perceptions of the usefulness of the RDCW system as a whole (recall that scores range from -2 to +2). The satisfaction component had a mean score of 0.6 (SD = 0.87), indicating a marginally positive feeling of satisfaction associated with the RDCW system. Additionally, usefulness and satisfaction scores in this analysis were significantly correlated ( $r = .719$ ,  $p < .001$ ), indicating that drivers who reported the RDCW to be useful were also likely to report the system to be satisfying (and vice versa).

*LDW.* Scale reliability tests for the LDW components of usefulness and satisfaction showed Cronbach's alphas of 0.93 and 0.88, respectively. The usefulness component for LDW had a mean score of 1.34 (SD = 0.78), while the satisfaction component had a mean score of 0.78 (SD = 0.87), indicating a marginally positive feeling of satisfaction associated with the LDW system. LDW usefulness and satisfaction scores were also significantly correlated ( $r = .746$ ,  $p < .001$ ).

*CSW.* Scale reliability tests for the CSW components of usefulness and satisfaction showed Cronbach's alphas of 0.96 and 0.93, respectively. The usefulness component for CSW had a mean score of 0.89 (SD = 1.09), while the satisfaction component had a mean score of 0.42 (SD = 1.1), indicating a marginally positive feeling of satisfaction associated with the CSW system. Similar to the scores for RDCW and LDW, CSW usefulness and satisfaction scores were significantly correlated ( $r = .807$ ,  $p < .001$ ).

### **Comparison across studies**

It is useful to compare these results to Van der Laan scores from a study of a different driver assistance technology. Doing so allows one to see how RDCW was perceived relative to other systems, and so may add a larger context within which to interpret the results from the present

study. The study under consideration is the automotive collision avoidance system (ACAS) FOT (Ervin et al., 2005). The ACAS FOT evaluated two different driver assistance systems: forward collision warning (FCW) and adaptive cruise control (ACC). The FCW system warned the driver of an emerging conflict that could lead to a rear-end crash via a combination of a forward radar and visual and auditory warnings as the driver approached another object. The ACC system was an enhancement of conventional cruise control that, in addition to controlling speed at a value selected by the driver (the set speed), also managed the distance to a preceding vehicle by automatically adjusting the vehicle’s throttle and brakes.

Ervin et al. (2005) was similar in many ways to the present study, in that both ACC and FCW were packaged together in one vehicle. In fact, the ACAS FOT was almost identical in design and method to the present study. Drivers in the ACAS FOT were randomly selected, licensed Michigan drivers. They drove an ACAS-equipped vehicle for 26 days, experienced the same system-disabled and enabled periods, and experienced almost identical orientation and debriefing procedures as in the RDCW FOT. In fact, the only substantial differences between the studies were the specific technologies being evaluated and slightly different sample sizes. This makes a comparison between the two studies particularly compelling.

Table 1 summarizes the mean Van der Laan scale scores for both studies. Notice that of all the technologies evaluated, ACC received the highest usefulness and acceptance scores. While this is interesting in itself, it is important to point out that ACC was not, first and foremost, a crash avoidance or warning feature, but rather a convenience feature. Thus, in terms of crash warning or mitigation, FCW has more in common with LDW and CSW. When comparing these Van der Laan scores, it can be seen that perceptions for CSW and FCW were very similar (i.e., both marginally positive). Additionally, the scores for LDW (and RDCW as a whole) were higher than those of CSW or FCW.

**Table 1. Mean Van der Laan scale scores from the RDCW FOT and ACAS FOT**

Subscale	RDCW FOT			ACAS FOT	
	RDCW	LDW	CSW	FCW	ACC
Usefulness	1.23	1.34	0.89	0.9	1.49
Satisfaction	0.6	0.78	0.42	0.5	1.48

Another interesting finding is the difference between the LDW usefulness and satisfaction scores, especially when compared to the same difference for ACC. For ACC, these two scores are almost identical. LDW, however, has a satisfaction score that is 0.56 lower than its usefulness score. This would suggest that while drivers found a similar level of utility with LDW and ACC, they were less satisfied with the LDW system. However it must be stressed that the same drivers did not rate both systems; each FOT had its own unique set of drivers.

## **RATINGS OF RDCW ALERT UTILITY**

As part of the debriefing, drivers viewed forward- and face-camera video for several of the alerts they received during weeks 2 through 4, and they were asked to provide a rating of usefulness for these alerts. The goal of presenting the video was to ascertain a rating of usefulness as it applied

to alerts in a variety of road scenarios, allowing drivers to “re-live” actual events they had experienced. Between 10 and 12 alerts ( $M = 10.9$ ,  $SD = 1.5$ ) were selected to be replayed to the returning driver (or fewer if the driver did not experience a particular type of warning).

All 78 drivers reviewed video and provided usefulness ratings for 870 alerts (441 for LDW and 429 for CSW). Twenty-two of these alerts were excluded from the following analyses because they occurred during trips that were later deemed invalid, leaving 848 alerts. In addition, ratings from drivers 1 and 2 were excluded because a different scale was used to measure their utility ratings (the protocol changed starting with driver 3). Consequently, an additional 23 alerts were excluded, leaving a total of 825 alerts (414 for LDW and 411 for CSW).

After having viewed a particular alert, the driver was asked whether the alert was useful (yes or no). The driver was also asked to rate how useful the alert was on a five-point Likert-type scale:

1	2	3	4	5
Not at all Useful	Slightly Useful	Somewhat Useful	Fairly Useful	Quite Useful

The drivers were asked to evaluate the alert not in terms of whether the alert was useful in a general sense, but rather to base their evaluation upon the specific driving situation and their behavior at the time of the alert. They were asked to recall, as much as possible, their state of mind and attentiveness to the driving task at the time when they received the alert. They were also instructed that *usefulness* could be defined as any quality that enhanced their driving experience or added some benefit to their driving. The drivers could review the alert as many times as they needed to. After providing the numerical rating, the drivers were asked to briefly explain in their own words why they provided the rating that they did for each alert.

Linear mixed-effects models were fit on the LDW and CSW utility ratings. In each case, a series of models was selectively compared to find the best fit. Each model initially included the dependent variable of utility rating and the independent variables of age group (three levels), gender, alert type (two levels: cautionary and imminent), the scenario associated with the alert, and whether the alert was associated with nondriving behaviors. Each model was then refit multiple times, each time excluding the main effect that was least significant. When only significant main effects remained, the model was refit again to include those main effects and their interaction terms. The nonsignificant interactions were removed to obtain the final model for each analysis. Each analysis also included random effects of “driver” and “driver by [within-subjects factor]” interactions. In other words, the random variance between drivers was included as a parameter within each model. Thus, if the effects of between-subjects variables (e.g., age or gender) or within-subjects variables (e.g., nondriving behaviors) were not significantly greater than random variance among drivers, they would not reach statistical significance in the model. Bonferroni corrections were used for all pairwise comparison tests.

When asked whether alerts were useful, drivers said that 74.9% of the LDW alerts were useful, compared to 54.3% of the CSW alerts (note that in both cases, the majority of alerts were rated as useful). When asked to give numerical utility ratings, drivers similarly rated LDW alerts as more useful than CSW alerts. The mean overall utility rating for LDW (collapsed across both cautionary and imminent alerts) was 3.3 ( $SD = 1.4$ ) while that of CSW alerts was 2.4 ( $SD = 1.4$ ).

Because the mixed-effects model analyses did not directly compare subsystems, it is unclear whether this difference is statistically significant.

The most notable differences between the LDW and CSW ratings were the higher frequency of *not at all useful* (1) ratings for CSW and the higher frequencies of *fairly useful* (4) and *quite useful* (5) ratings for LDW. Stated another way, just over 40% of the CSW alerts reviewed were rated as *not at all useful*, while about 50% of the LDW alerts reviewed were rated as *fairly useful* or *quite useful*.

To further parse these results, the mean utility ratings were broken down by cautionary and imminent alerts for each system (Figure 1). The error bars in the figure represent the standard error of the mean. While the mean utility ratings for cautionary and imminent LDW alerts were roughly equivalent, the mean rating for imminent CSW alerts was lower than for cautionary CSW alerts, a finding that was statistically significant,  $F(1, 338) = 5.3, p < .05$ . While 50% of the imminent CSW alerts were rated as *not at all useful* (1), only 33.2% of the cautionary CSW alerts received the same rating. In contrast, 15.2% of the cautionary alerts were rated as *quite useful* (5), while only 9.5% of the imminent alerts received that rating.

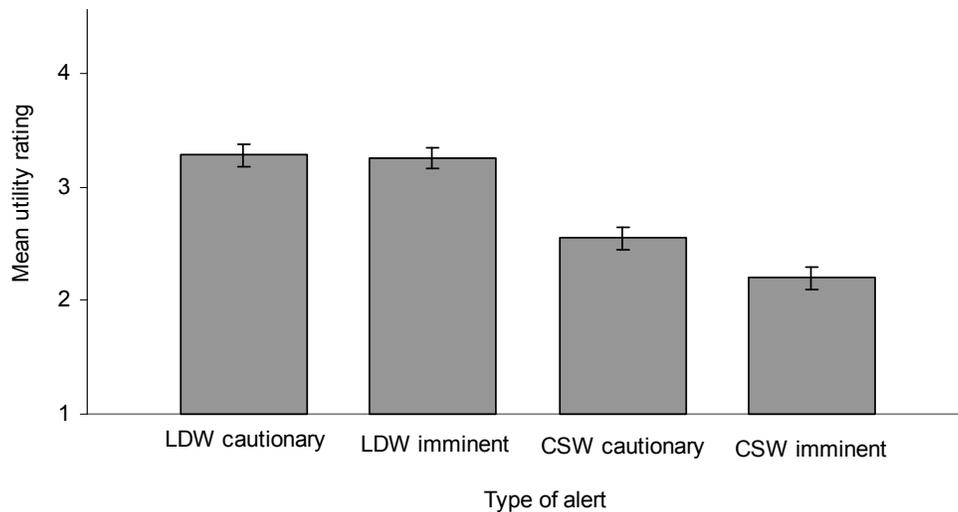


Figure 1. Mean ratings of utility for cautionary and imminent LDW and CSW alerts

## CONCLUSIONS

In sum, while drivers had generally positive perceptions of the RDCW system, including each individual subsystem, they found LDW more useful and satisfying than CSW. Thus, as might be expected, the scores for the entire RDCW system fell in between those of LDW and CSW. Drivers generally gave higher ratings for usefulness than they did for satisfaction. This was true for the entire system, as well as for LDW and CSW individually. This supports the findings from the focus groups and the rest of the post-drive questionnaires in which drivers saw a functional or conceptual value to RDCW, but were not completely satisfied with some aspects of the system.

Most drivers reported that the RDCW system would increase driving safety and that they would be somewhat willing to purchase the RDCW system. While LDW and CSW systems, and the

combined RDCW system, were generally received positively, most drivers made distinctions between the LDW and CSW systems in terms of their perceived utility. Ratings of utility of individual alert events by the drivers found 75 percent of LDW and 54 percent of a sample of the CSW alerts to be useful. Drivers reported a willingness to purchase LDW more often than they reported the same willingness in regard to CSW. LDW was considered by a number of drivers to be a generally desirable feature. CSW was also seen as a useful concept and was frequently described as being useful in unfamiliar surroundings or poor weather. Drivers subjectively reported that they were better drivers when LDW was enabled, using cell phones less often and turn signals more often, and drivers infrequently reported concerns regarding false alerts from LDW. The companion paper at this conference addresses the system description, experimental design, performance of the road departure system in naturalistic use, and analyses of safety impacts of the technology

## ACKNOWLEDGMENTS

The authors thank colleagues at Visteon (Debby Bezzina, Tim Tiernan, Faroog Ibrahim, Dinu Madau, and others) and AssistWare/Cognex Corporation (Dean Pomerleau). Nissan Corporation graciously provided engineering information to facilitate system integration into the vehicle. The work was conducted under agreement DTFH61-01-X-00053 with the U.S. DOT. Our thanks to Lloyd Emery (retired) and Ray Resendes of NHTSA and Bruce Wilson of the Volpe Center.

## REFERENCES

- Emery, L., Srinivasan, G., Bezzina, D., LeBlanc, D., Sayer, J., Bogard, S., and Pomerleau, D. (2005). Status report on USDOT project "An intelligent vehicle initiative road departure crash warning field operational test." *Proceedings from the 19<sup>th</sup> International Technical Conference on the Enhanced Safety of Vehicles* (US DOT HS 809 825). June 2005. Washington, DC: Department of Transportation.
- Ervin, R., Sayer, J., LeBlanc, D., Bogard, S., Mefford, M., Hagan, M., Bareket, Z., and Winkler, C. (2005). Automotive collision avoidance system (ACAS) field operational test methodology and results (US DOT HS 809 901). Washington, DC: Department of Transportation.
- LeBlanc, D., Sayer, J., Winkler, C., Bogard, S., Devonshire, J. Mefford, M., Hagan, M., Bareket, Z., Goodsell, R., and Gordon, T. (2006). Road departure crash warning system (RDCW) field operational test final report (Report No. UMTRI-2006-9-1). Ann Arbor: University of Michigan Transportation Research Institute.
- LeBlanc, D., Sayer, J., Winkler, C., Bogard, S., and Devonshire, J. (2007). Field test results of a road departure crash warning system: Driver utilization and safety implications. *Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training, & Vehicle Design*. Stevenson, WA.
- Van Der Laan, J.D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research*, 5(1), 1-10.