

Developing Complex Crash Warning Simulations for Human Factors Evaluations

Paul Green
UMTRI-Human Factors Division
University of Michigan
Ann Arbor, Michigan 48109-2150 USA

In traditional vehicle warning experiments, each subject sees each warning once, and warning comparisons are between subjects, a very inefficient approach. Driving simulator experiments for the recent RDCW and IVBSS projects used within-subjects designs, with each subject responding to a warning about once per minute. That presentation rate and the scenarios developed seemed reasonable to independent human factors experts who experienced them and the simulations were often able to distinguish differences of interest.

What made the simulations reasonable was (1) the use of real world crash data as a starting point for scenarios, (2) the large number of scenarios employed, each of which had both normal and crash-related outcomes, (3) the presence of three to five vehicles in the scene, each requiring the driver's attention, (4) the use of real on-road data to select gaps, closing rates, etc., and (5) work-arounds when scenarios did not go according to plan. For developing and communicating the plan, option tables and storyboards were particularly useful.

INTRODUCTION

In the United States, about 42,000 people are killed each year in motor vehicle crashes (U.S. Department of Transportation, 2006) and, depending on how an injury is defined, the annual number of injuries sustained in crashes is 50 to 100 times that amount. For the last few decades, motor vehicle engineering programs have emphasized passive safety/vehicle crashworthiness – energy-absorbing vehicle bodies, seat belts, child restraint systems, and air bags. However, the physical limitations of how much energy vehicle structures can absorb is being approached, so the emphasis is now on active safety/crash avoidance, especially vehicle stability control and crash warning systems. To avoid exposing subjects and the public to unproven warning systems in life-threatening situations, driving simulators are often used prior to on-road testing to examine warning system effectiveness.

PROBLEMS TO OVERCOME IN ALL DRIVING SIMULATOR WARNING EXPERIMENTS

In most projects, the funding and schedule for testing alternative warnings is limited. Usually that means conducting only a few experiments and about a month for data collection for each experiment. For the sake of discussion, assume one has 6 weeks for data collection for each experiment and the simulator operates on one shift. One would therefore have 240 hours available for testing, of which half is for pilot tests and subsequent modifications, so only 120 hours are available for “real” testing. With no shows, gaps between subjects, etc., the actual number might be 60 hours. What can one learn about real driving in 60 subject hours (say 30 subjects for one two-hour session each), a moderate size experiment?

Problem 1: If crashes are to occur with absolute real-world frequency, one cannot collect enough data within the timeframe of a typical experiment.

According to the National Household Travel Survey (U.S. Department of Transportation, 2003), the average person in the U.S. travels 14,500 mi/yr, with about 90% by a personal motor vehicle (about 13,000 mi/yr). Assuming a mean speed of 40 mi/hr yields 325 hours of driving time per year. The average driver is involved in a crash about once every five years, or about once every 1,625 hours of driving.

However, in 60 subject hours, given real world crash frequencies, the chance of even one crash in an experiment is low (by more than an order of magnitude) and, for statistical analysis, more than one crash is desired. Fortunately, the concept of accelerated testing is well accepted in engineering and science, which is consistent with greater exposure to crash provocative events (and many crashes occurring). Nonetheless, an unnaturally high frequency of crashes is likely to be the first criticism of driving simulator experiments.

Problem 2: The traditional one-crash-warning-per-subject experiment does not provide enough data for the effort required.

As an example of a traditional warning experiment, in Yoo, Hunter, and Green (1996), subjects drove a simulated two-lane road. Subjects passed 17-parked cars and briefly followed two cars that pulled out from the side of the road (well ahead of the subject) and then sped off. At about 15 minutes into the simulation, a car that had been parked on the side of the road pulled out in front of the subject, forcing them to brake or swerve to avoid a collision.

The test sample consisted of three groups of eight subjects, with each group receiving a different warning. There were five crashes in response to the “swerve left” text

warning, three for the icon warning, and seven for the group without a warning. Some 13 minutes later, the pull out scenario was repeated, but there were only two crashes (one for text, one for no warning).

After the first crash/near crash event, subjects became extremely vigilant for a repeat of that scenario, slowing when there was a car on the side of the road. Consequently, in this type of experiment, each subject should respond only once to each warning, and experiments must be designed for analysis between subjects, with all the consequent problems related to large sample sizes and matching groups.

AN APPROACH DEVELOPED IN THE RDCW AND IVBSS PROJECTS

Over the last few years, UMTRI has completed the Road Departure Crash Warning (RDCW, Sayer, Cullinane, Zylstra, Green, and Devonshire, 2003) project and phase 1 of the Integrated Vehicle-Based Safety System (IVBSS, Green, Sullivan, Tsimhoni, Oberholtzer, Buonarosa, Devonshire, Schweitzer, Baragar, and Sayer, 2008) project. RDCW involved two warning systems, one for leaving the road (lateral drift or lane departure warning, LDW) and a second for approaching a curve too fast (curve speed warning, CSW). In IVBSS, forward crash warning (FCW) and lane change-merge (LCM) warnings were added, as well as blind spot detection (BSD), so in some sense there were five systems.

As these projects involved multiple warnings, not single warnings, the issues of warning confusion and response to concurrent warnings were important. Further, the government mandated project schedules were extremely tight, so many questions needed to be answered using a limited number of subject hours, and in a very short timeframe. Some questions, especially those relating to the test protocol, had to be addressed based on expert opinion. Unfortunately, this situation of limited resources will become the norm in the future.

How should the plans for experiments be developed?

For coordination in IVBSS, written plans were developed for four major simulator experiments (and several smaller ones). What evolved was a 33-page internal document that was incorporated in parts into the final report (Green, et al., 2008). The internal document contained (1) the research questions, (2) storyboards for all scenarios (described later in this paper), (3) experimental task duration estimates, (4) sketches of the test world and details of the counterbalancing, (5) how each warning was presented, including false alarms, (6) the reasoning for experimental constraints and requirements (why the world was a loop, how events would be triggered, etc.) and (7) how subjects could be induced to execute the desired maneuvers. Although writing the document was time-consuming, it facilitated decision-making and served as a useful reminder when questions arose about the test methods late in the project (long after the reasons would normally be forgotten, such as when the final report was written). Such a document should not be included

as a contract deliverable, because too much time will be spent on the appearance of the document and undesired project micromanagement could occur. However, it could prove particularly useful for dissertation research.

How should the crash scenario options be determined?

Planning of crash scenarios should begin with a review of the relevant crash literature, which is sometimes neglected by those new to driving simulator research. That literature provides typologies used to describe crash types (e.g., Massie, Campbell, and Blower, 1993; Najm, Smith, and Yanagisawa, 2007) as well current data on the frequency of particular crash configurations. For example, Najm, Smith, and Smith (2001) reported that the most common crossing path crash (30% of all crossing path crashes) was straight crossing path (SCP, “T-boned,” Figure 1), with the left turn across path/opposite direction configuration (LTAP/OD) in which someone making a left turn is struck by an oncoming vehicle being second most common (28% of all crossing path crashes). Thus, a study of crossing path crashes should simulate these crash configurations.



Figure 1. Common intersection crash configurations
Source: Najm, Smith, and Smith, 2001, p. 6.

In IVBSS, there were eight scenarios ((1) lead vehicle brakes, (2) lead vehicle brakes after the subject changes lanes, (3) reveal (described later), (4) cut in, (5) merge from blind spot, (6) accelerate and merge from blind spot, (7) wind gust, (8) excessive curve speed). To develop these scenarios and assure all options were considered, tables were created listing all subject maneuvering options in response to the actions of other vehicles. Table 1 presents examples for FCW maneuvers. A good experiment should cover as many positioning and maneuvering combinations as possible, and for every combination, there should be crash and non-crash outcomes.

In IVBSS, such a table highlighted the need for reveal scenarios in which the lead vehicle abruptly changes lanes to reveal a stopped car ahead of it (“the reveal scenario”), and the subject has little time to react (especially if the stopped car is small and similar in color to the road surface). Ordinarily, when a lead vehicle changes lanes, the subject has more maneuvering space, and their eyes often follow the lead vehicle, so in a reveal situation, they often miss the parked car in their lane. A surprisingly large number of real-world crashes involve stopped or slowly moving vehicles in the driver’s lane, even on expressways.

Table 1. Possible Subject Actions: FCW Vehicle Position and Maneuver Combinations

		Vehicle Ahead (Lead Vehicle)								
		Starts	In Left Lane		In Right Lane		Changes to Left Lane		Changes to Right Lane	
		Then	Brakes	Continues	Brakes	Continues	Brakes	Continues	Brakes	Continues
Subject	In left lane	brake or swerve	continue	continue or swerve	continue	brake (cut in)	continue	swerve or brake	continue (if reveal, brake)	
	In right lane	continue or swerve	continue	brake or swerve	continue	swerve or brake	continue (if reveal, brake)	brake (cut in)	continue	
	Changes to left	brake	continue	swerve or continue	continue	brake	continue	swerve	continue	
	Changes to right	swerve or continue	continue	brake	continue	swerve	continue	brake	continue	

Except for those related to blind spots, maneuvers involving lead vehicles and lane changes are similar, regardless of whether the subject is in the left or right lane. This similarity provides some flexibility in designing experiments, as situations can be repeated without being identical.

For this idea to work, some control over the lane choice is needed. To control choice, cars adjacent to the subject blocked lane changes or construction barrels forced lane changes. In combination, this allowed for creating some very hazardous merges.

How does one keep subjects from anticipating which scenarios will lead to a warning?

In IVBSS, the variety of driving situations made warnings unexpected. There were four tables of conditions (Table 1), though there were few conditions for LDW because that warning was usually induced by a lateral wind gust (with three levels of intensity) which subjects clearly could not see (or in IVBSS, hear, a weakness). Furthermore, there were always three (sometimes five) vehicles visible, all of which were moving and required attention from the subject. Finally, about 1/3 of all planned warnings were false alarms, which is reasonably close to the performance of some real-world warning systems (Ervin, Sayer, LeBlanc, Bogard, Mefford, Hagan, Bareket, and Winkler, 2005). Thus, when a warning was presented, the subject should not respond blindly but needed to check the traffic situation.

To simplify the option space, consider minimizing the number of lanes of traffic.

With more lanes, there are more maneuvering options that need to be considered (and potentially counterbalanced), though it does make the simulation more interesting. For example, with three travel lanes in each direction, a subject in the center lane would need to simultaneously consider vehicles ahead, to the left, and to the right (as well as behind). However, the greater the number of lanes, the greater the number of vehicles that need to be programmed into the simulation, increasing software development time. Of course, when selecting scenarios, real-world crash data must be considered, so if a crash type of interest often occurs on a

certain road type, then that road type should be simulated. However, be prepared for the programming effort.

How should the crash-provoking scenarios be communicated to others?

Communication is important both within the research team and between the team and the sponsor. In IVBSS, storyboards (Figure 2) were created to show sequences of actions. Storyboards were originally developed to plan animated cartoons, but are now used in movie production, video game design, and human-computer interface design (Snyder, 2003). Storyboards helped establish vehicle placement and gaps, with speeds and accelerations being added as a second step. The sequences of storyboards helped communicate timing.

Cultural differences should be considered when creating storyboards. In the U.S., emphasis would be placed on getting the graphics to resemble real vehicles. In Japan, the vehicles would be cartoons, which take less time to create.

One of the more interesting aids to planning scenarios was a set of Matchbox toy cars (purchased as “miniature research vehicles,”) that could be moved around a desktop (a simulated road). Sometimes, experimenters got carried away in making sounds when simulating crash scenarios.

How often should crash-provoking scenarios occur and what are the implications for experiment design?

In the RDCW and IVBSS projects, crash-provoking scenarios were planned to occur about once per minute based on the following reasoning. After two hours, subjects get bored, and performance declines and becomes more variable. For a two-hour session, at least 40-45 minutes are required for pre-test activities (forms, instructions, practice driving, and other tasks) and post-test activities (survey, payment), leaving 75-80 minutes for actual testing. To minimize sickness, test blocks are typically 10 to 15 minutes long with some time between blocks to save the data, check for motion discomfort, and give subjects a short break. A typical experiment involves four to six test blocks, though there could be as many as eight blocks.

In simulator experiments, the implementation of the warning system is typically a within-block factor, constant for

a subject over a series of trials. For 4-6 blocks, one could use a 2x2 or 2x3 design (e.g., two modality combinations such as auditory alone and auditory plus visual combined with, say, two or three intensity levels (subtle, normal, loud) for the

auditory warning). Within blocks, there is time for 10-14 events that lead to warnings. To reset the subject's mental state at the beginning of each block, there are often no crash-provoking events in the first minute or so.

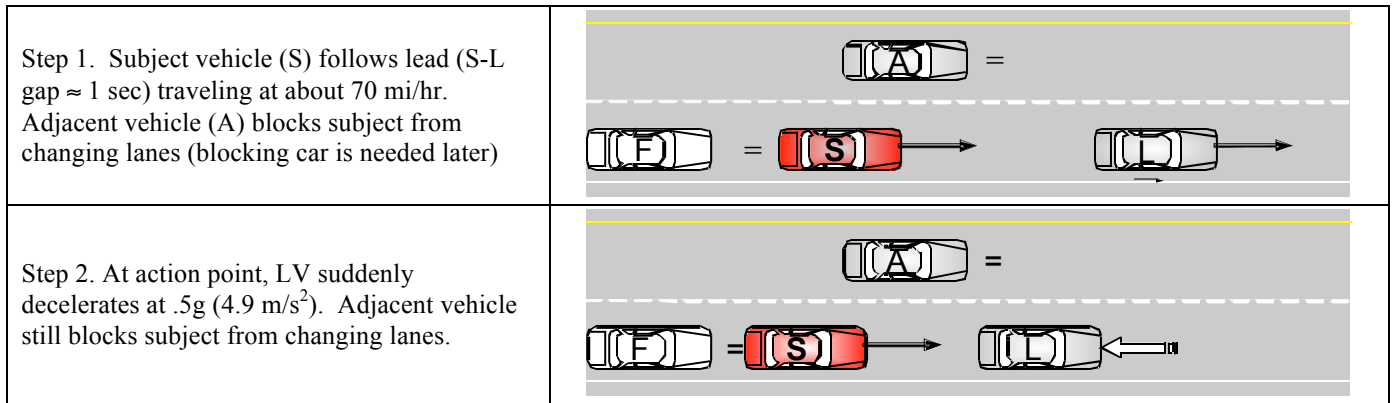


Figure 2. Two of the Four Steps from IVBSS Scenario 1 Storyboard, FCW: Lead Vehicle Suddenly Decelerates

How much time should pass between crash-provoking events?

In the worst case, a subject's initial speed is zero (at the beginning of a block or stopping after a crash). Therefore, time is needed to (1) accelerate to the posted speed, (2) drive beyond what they can see after they have reached the posted speed, and (3) establish a stable situation. A contemporary car can reach expressways speeds of 65 to 70 mi/hr (95 to 103 ft/s) in less than 10 seconds if the throttle is fully open. For a more moderate acceleration, assume double that time period, 20 seconds. In daylight scenes, simulators will easily show an additional 1,000 feet or more of scene ahead of the vehicle, depending on the cutting plane. It takes about 10 seconds to drive 1,000 feet. But, if something happens as soon as one reaches that point (30 seconds after start), subjects will anticipate the event, so a factor of two safety margin suggests a duration of about one minute between events, with some variability to make crash-provoking events unpredictable. This rationale led to events occurring at roughly once per minute (\pm 10 s, roughly) in the RDCW and IVBSS studies.

How can scenarios be made to appear realistic?

The motion of vehicles, the difference in speed of vehicles in traffic, acceleration rates, braking rates, etc. should be reasonable for real vehicles. In UMTRI projects, the results in the Automotive Collision Avoidance System (ACAS) (Ervin, Sayer, LeBlanc, Bogard, Mefford, Hagan, Bareket, and Winkler, 2005) and RDCW (LeBlanc, Sayer, Winkler, Bogard, Devonshire, Mefford, Hagan, Bareket, Goodsell, and Gordon, 2006) reports provided real-world data for estimating the value of these variables, which was supplemented by comments and caveats provided by the vehicle dynamics experts at UMTRI.

What should happen if a crash occurs?

In the real world, when there is a crash, things come to a halt. If that occurs in a simulator, one would only have partial data for that subject, making analysis more difficult. One could replace all subjects that crash so there is no missing data, but subjects who crash are of great interest, so throwing their data away is unwise. But, a simulator is not the real world. In UMTRI studies, making the subject "invincible" proved to be useful, so that trials were not terminated when contact occurred. For the DriveSafety simulator used, subjects were able to drive through other vehicles in the world. (Subjects were not informed of this in advance, so they tried to avoid crashes as in real driving.)

Another strategy considered (which is difficult to implement but possible in a virtual world) is to change the laws of physics. Thus, if a subject was just about to crash into a lead vehicle, once it had been determined a crash would ordinarily occur, that lead vehicle would suddenly accelerate at a rate an ordinary motor vehicle could not achieve (say 0 to 60 mi/hr in 1 second), so there would not be contact. Unfortunately, crashes into fixed objects are still unavoidable.

What happens if the vehicle configuration is not as desired?

In the IVBSS project, the scenarios were quite complex relative to previous experiments, and given the tight schedule, it was not possible to explore all possible subject responses and create contingencies for them. For example, in the IVBSS forced lane change scenario, the plan was for subjects to maneuver among vehicles, but sometimes they merged behind the group. Ideally, the simulation should have incorporated a pause function that would freeze the traffic and allow the subject to drive (or be warped) to the desired location, after which the simulation restarted. For IVBSS, there was not time to develop and test this feature, so the few subjects for whom this occurred were sometimes instructed to drive through

objects in the simulation (i.e., intentionally crash). Admittedly, this work-around was awkward, but the alternative was to replace the data from any subject that lost position at any time. Future studies will include the freeze feature.

What happens when warnings are not triggered as planned?

Table 2 shows the frequency of occurrence of the four warnings in the first major simulator experiment (experiment 2 in the final report). “Triggered not as planned” warnings occurred because subjects did something unexpected—they

drifted out of the lane when responding to an FCW, they changed lanes at the wrong time and got too close to a lead vehicle, etc. Likewise, “planned but not triggered” situations also occurred when subjects behaved in unexpected ways—they drove too slowly/distant from a lead vehicle to trigger an FCW, they quickly recovered from a wind gust, etc. For IVBSS, there were so many useful responses to warnings that it was not worth the months of effort (nor was there time in the project schedule) to track down why each scenario did not always go as planned and then write (and exhaustively test) software to eliminate departures from the plan. For some analyses the “not as planned” data was not used. Other times, the “not as planned” data was filtered and some of it was used.

Table 2. Frequency of Warnings in the IVBSS Experiment

Warning Category	FCW (658)		CSW (225)		LDW (613)		LCM (159)	
	Real	FA	Real	FA	Real	FA	Real	FA
Triggered as planned	280	81	62	34	88	14	49	61
Triggered not as planned	282	15	129	0	510	1	48	1
Planned but not triggered	9	1	1	0	16	16	79	4
Total triggered (1655)	562	96	191	34	598	15	97	62

CLOSING THOUGHTS

Multiple warning systems with complex interactions developed on tight schedules are now the norm, as opposed to the single warning systems of the past. To develop these systems, rethinking about how studies are planned is needed, with the RDCW and IVBSS being quite insightful. Perhaps the most important finding reported was that if there are enough warnings and scenario variety, warnings could reasonably be presented quite often (here, once per minute). Unfortunately, there are no experimental data comparing this rate of presentation with others. Nevertheless, all of the human factors experts (both UMTRI staff and independent, external evaluators) who drove the RDCW and IVBSS simulations felt the rates were reasonable and the simulations were able to distinguish some of the differences of interest.

REFERENCES

Ervin, R., Sayer, J., LeBlanc, D., Bogard, S., Mefford, M.L., Hagan, M., Bareket, Z., and Winkler, C. (2005). Automotive Collision Avoidance System (ACAS) Field Operational Test – Methodology and Results (Technical Report HS 809 901), Washington, D.C.: U.S. Department of Transportation.

Green, P.A., Sullivan, J.M., Tsimhoni, O., Oberholtzer, J., Buonarosa, M.L., Devonshire, J., Schweitzer, J., Baragar, E., and Sayer, J.R. (2008). Integrated Vehicle-Based Safety Systems (IVBSS): Human Factors and Driver-Vehicle Interface Summary Report (Technical Report DOT HS 810 905), Washington, D.C.: National Highway Traffic Safety Administration, U.S. Department of Transportation (<http://www.itsa.org/itsa/files/pdf/IVBSS2008HFDVI.pdf>).

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 (UMTRI Technical Report 2006-9-1), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Massie, D.L., Campbell, K.L., and Blower, D.F. (1993). Development of a Collision Typology for Evaluation of Collision Avoidance Strategies, Accident Analysis and Prevention, 25(3), 241 -257.

Najm, W.G., Smith, J.D. and Yanagisawa, M. (2007). Pre-Crash Scenario Typology for Crash Avoidance Research (Technical Report DOT HS 810 767), Washington, D.C.: U.S. Department of Transportation, National Highway Traffic Safety Administration.

Sayer, J., Cullinane, B., Zylstra, B., Green, P., and Devonshire, J. (2003). Lateral Drift and Curve Speed Warnings: A Driving Simulator Evaluation of Auditory and Haptic Implementations (UMTRI Technical Report 2003-41), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Snyder, C. (2003). Paper Prototyping. San Francisco, CA: Morgan Kaufmann.

U.S. Department of Transportation (2003). Highlights of the 2001 National Household Travel Survey (Technical Report BTS03-05), Washington, D.C.: U.S. Department of Transportation, Bureau of Transportation Statistics.

U.S. Department of Transportation (2006). Traffic Safety Facts 2005 (Technical Report DOT HS 810 631), Washington, D.C.: U.S. Department of Transportation, National Highway Traffic Safety Administration, National Center for Statistics and Analysis.

Yoo, H., Hunter, D., and Green (1996). Automotive Collision Warning Effectiveness: A Simulator Comparison of Text vs. Icons (UMTRI Technical Report 96-29), Ann Arbor, MI: University of Michigan Transportation Research Institute.