2004-01-0450

Comprehension Testing of Active Safety Symbols

John L. Campbell
Battelle Human Factors Transportation Center

David H. Hoffmeister
Ford Motor Company

Raymond J. Kiefer
General Motors Structure and Safety Integration Center

Daniel J. Selke
Mercedes-Benz USA, LLC

Paul Green
University of Michigan Transportation Research Institute

Joel B. Richman
Battelle Human Factors Transportation Center

ABSTRACT

This paper describes an effort to develop a valid and reliable process for comprehension testing of candidate automotive symbols and to conduct comprehension testing on a set of new symbols being considered for in-vehicle active safety systems. The comprehension testing process was developed though a multi-year effort, supported by Society of Automotive Engineering (SAE) and other organizations, aimed at generating a test methodology that would: yield high-quality comprehension data for new automotive symbols, provide clear and specific guidance back to symbol developers based on the test results, and could be adopted and performed internationally to support international standards efforts. Seventeen (17) candidate symbols were evaluated for three classes of in-vehicle active safety systems: forward collision warning (4 symbols), side collision warning (6 symbols), and lane departure warning (7 symbols). So far, testing has been completed in Germany, Sweden, Japan, and the United States. In the US testing, the study yielded comprehension data, appropriateness rankings, and diagnostic design feedback for all 17 icons tested. Based on these data, US recommendations have been made to the International Standards Organization (ISO) for all three classes of in-vehicle safety systems. This paper describes the process associated with developing the procedure, including the international outreach required to obtain support from major ISO countries, as well as the methods and results from the US testing.

INTRODUCTION

Icons can be used to communicate information to the driver in a language-free and space-efficient manner (Green, 1993). Incomprehensible icons have the potential to affect safety (e.g., if the driver does not understand the icon). However, despite the ubiquity of icons and symbols within the in-vehicle environment, few guidelines exist for testing of icons. Key shortcomings of existing icon testing procedures (e.g., ISO/DIS 9186, 1988; Wolff & Wogalter, 1998) include: a lack of contextual information provided to experimental subjects, over-reliance on evaluator judgment regarding how well an experimental subject comprehended an icon's meaning, and a lack of prescriptive information – feedback to icons designers regarding how individual icons could be improved based on the results of comprehension testing (Richman, Campbell, & McCallum, 2002).

Under contract to the Federal Highway Administration (FHWA), the Battelle Human Factors Transportation Center developed a set of design guidelines (Campbell, Richman, Carney, & Lee, 2002) that included an integrated set of guidelines for evaluating in-vehicle icons. In the fall of 2001, a Symbols Working Group was formed from within the SAE Safety and Human Factors Committee to improve upon and extend these guidelines. The group (all of the authors of this paper except for J. Richman) was comprised of human factors experts from the transportation research and automotive industry. The groups' stated objective was to develop and test a general process to evaluate icon/symbol comprehension that could be used internationally and –

eventually – generate active safety symbols that could be incorporated into SAE J2402 (ISO 2575).

After several months of interactions within the Symbols Working Group, a revised testing procedure was agreed upon and then presented to ISO member countries during an ISO meeting in Portland, Oregon in June, 2002 of ISO Technical Committee 22, Subcommittee 13, Working Group 5 (Ergonomics of Road Vehicles, Symbols). The test procedure was well-received at this meeting, and in late 2002 and early 2003, icon comprehension testing using the new procedure was conducted by Sweden, Germany, and Japan. In these tests, 17 candidate symbols were evaluated for 3 classes of in-vehicle safety systems: forward collision warning (4 symbols), side collision warning (6 symbols), and lane departure warning (7 symbols).

Eleven of the 17 candidate active safety symbols chosen for evaluation were drawn from symbols developed as part of a rank order testing study conducted by Balbale and Kiefer (2001), which included participants from the United States, Japan, and Sweden. The goal of this earlier study, and an underlying goal of the current work, is to develop an integrated set of active safety symbols which facilitate driver comprehension from an international perspective. The additional six candidate symbols evaluated in the current work were suggested from Working Group members (and the organizations they represented) and ISO member countries.

A summary follows describing the conduct and results of a study to evaluate driver comprehension of the 17 candidate symbols in the United States. In the current testing, three standard icons were also included in the testing in order to provide comparison data for the other 17 candidate symbols.

METHOD

Seventy-seven (77) subjects from the Seattle, Washington metro area participated in the icon comprehension testing. They were recruited through newspaper ads; flyers posted on a university campus, in local establishments, libraries, and other community gathering places; and by word-of-mouth. Subjects initially contacted Battelle to express interest in participation. This initial telephone screening made sure that all potential subjects: (1) had an active driver's license, (2) had at least two years of driving experience, (3) were over 18 years of age, and (4) matched desired combinations of age and gender. We wanted to have roughly equivalent proportions of male and female subjects, as well as roughly equivalent proportions of younger, middle-aged, and older subjects. None of the subjects were employed in the automotive design or engineering industries. At the conclusion of this initial telephone screening, individual subjects were scheduled for a data collection session. The 77 subjects were tested across 3 days, during which 14 data collection sessions were conducted. Some sessions had as many as 14 subjects, others had as few as 3. Table 1 shows the age, gender, and number of participants in the study.

Within each of the three groups of candidate symbols, there were symbols that contained the same elements and looked similar to one another. To avoid learning effects among the subjects, different candidate symbols for the same message were separated into two different test sets/booklets. Thus, each subject only evaluated ten (10) distinct candidate icons (8 or 9 of the candidate icons, plus 1 or 2 of the standard icons). For each booklet, the order in which subjects evaluated the individual icons was varied, resulting in five distinct "orders" of the icons for each of the two booklets.

Table 1. Age, gender, and number of participants in the U.S. study.

Age group	No. of participants	Participant Gender			
		Male		Female	
		Booklet 1	Booklet 2	Booklet 1	Booklet 2
18-25	17	2	2	4	9
26-40	22	5	4	7	6
41-54	24	10	3	6	5
55+	14	2	5	3	4
Σ	**77**	**19**	**14**	**20**	**24**
		33		**44**	

During the testing, subjects were provided with general descriptions of the in-vehicle technologies associated with the symbols, to provide some context to the testing. The introduction to the testing provided in each subject's response booklet – and read aloud to the subjects – is shown below.

Testing consisted simply of providing this context as part of the introduction to a subject response booklet and asking subjects – for each icon/symbol – "What do you think that this symbol means?" – as shown below in Figure 1. Subjects wrote their response to this question for each icon (each icon was presented on a separate page in the response booklet), and then moved on to the next icon.

INTRODUCTION

Thank you for participating in this study. Our purpose is to investigate issues related to the use of graphical symbols or icons in the in-vehicle environment. Recent advances in automotive technology have allowed the development of various devices that can present a wide range of safety and vehicle information to drivers. Much of this information will be provided to drivers through color displays located on the instrument cluster or center console of the vehicle. These display systems might present information such as:

Vehicle Condition Monitoring: *Device that informs the driver of current status or any problems with the vehicle or vehicle systems. For example, information presented to the driver could reflect information or warnings about fuel levels, oil pressure, tire pressure, engine temperature, seat belts, or air bag systems.*

Trip Navigation: *Device that provides the driver with route assistance and trip planning information. For example, this could include driving directions, cautions about certain routes, current traffic conditions, road construction, or weather updates.*

Motorist Services: *Device that provides an in-vehicle "Yellow Pages" function for drivers. For example, this directory could present information regarding the availability and locations of motorist services such as gas stations, restaurants, hotels, or recreational activities.*

Collision Avoidance System: *On-board sensor and display systems that detect on-coming vehicles or other unsafe driving conditions and warn the driver of an impending collision. For example, these collision warnings could reflect a possible collision with another vehicle or object, a vehicle wandering out of its lane, or other unsafe traffic situations.*

You are driving in your car and you suddenly notice the following yellow or red indicator on your dashboard light up:

What do you think this symbol means?

Figure 1. Standard situation/question presented to the subjects for each icon tested.

After the comprehension portion of the study was complete, subjects were asked to rank order all of the candidate icons according to how well they thought that the candidate icons represented the message that they were intended to convey. For the ranking portion of the testing, subjects were given the following written description for the forward collision warning symbols: *"Warning: You may be about to crash into a car in front of you;"* the following written description for the side obstacle warning symbols: *"Warning: There is a vehicle or other obstacle next to your car;"* and the following written description for the lane departure warning symbols*: "Warning: Your vehicle is about to drive outside the driving lane."* It should be noted that the comprehension portion of this testing is considered to be of more direct relevance in assessing driver comprehension than this rank order testing, which can be used to help distinguish between icons with similar comprehension scores. However, rank order testing does provide an efficient means of downsizing a group of candidate icons for subsequent comprehension testing (as was done with the Balbale and Kiefer [2001] results), which is relatively more time-consuming and resource intensive.

The descriptions shown in Table 2 were used for categorizing the subject's responses into comprehension scores (i.e., to make the comparison between perceived and intended meaning).

RESULTS

Table 3 provides the key results from this study. Looking at the table from left-to-right, the first column shows the individual icons, the icon number, and whether the icon was presented in Booklet 1 or Booklet 2. The next nine columns provide the raw scores – in percentages – corresponding to the nine scoring categories described in Table 2. The next four columns provide combined percentages for key categories. The first column, labeled "1-2 High," is defined as high comprehension resulting from the combined scores from scoring categories 1 and 2. The column labeled "3-4 Low" reflects the combined scores from scoring categories 3 and 4 and is defined as low comprehension. The column labeled "5-8 None" reflects the combined scores from scoring categories 5, 6, 7 and 8 and is defined as no comprehension. The column labeled "9 Crit. Con." reflects the scores from scoring category 9 and denotes a critical confusion. Critical confusions or errors reflect responses that indicate that the subject perceived the message to convey a potentially unsafe action. For example, a forward collision warning being perceived by a subject to indicate a rear collision, or a lane departure warning being perceived by a subject to indicate that they should steer the vehicle off the side of the road.

The table off-set to the right in Table 3 shows ranking data for each icon; i.e., percentage of participants that believed that the icon was the most appropriate symbol for a particular message (rank of 1), versus the second most appropriate symbol for a particular message (rank of 2), etc.

Table 2. Rating scales for categorizing and scoring subject responses to the icons.

Comprehension Score	Description
1	The response matches the intended meaning of the icon exactly.
2	The response captures all major informational elements of the intended meaning of the icon, but is missing one or more minor informational elements.
3	The response captures some of the intended meaning of the icon, but it is missing one or more major informational elements.
4	The response does not match the intended meaning of the icon, but it captures some major or minor informational elements.
5	The response does not match the intended meaning of the icon, but it is somewhat relevant.
6	Participant's response is in no way relevant to the intended meaning of the icon.
7	Participant indicated he/she did not understand the icon.
8	No answer.
9	For safety-critical icons, identify the number and percentage of critical confusions or errors. Critical confusions or errors reflect responses that indicate that the subject perceived the message to convey a potentially unsafe action.

Table 3. Icon comprehension ratings and rankings.

Icon Comprehension Ratings

Icon*	1	2	3	4	5	6	7	8	9	1-2 High	3-4 Low	5-8 None	9 Crit Con
1 1	8%	15%	13%	3%	23%	3%	8%	0%	28%	23%	15%	33%	28%
2 1	8%	26%	46%	3%	5%	10%	3%	0%	0%	33%	49%	18%	0%
3 2	13%	32%	16%	26%	5%	0%	0%	3%	5%	45%	42%	8%	5%
4 2	3%	8%	5%	3%	13%	55%	13%	0%	0%	11%	8%	82%	0%
5 1	10%	23%	5%	3%	33%	8%	13%	0%	5%	33%	8%	54%	5%
6 2	34%	16%	5%	8%	8%	0%	13%	0%	16%	50%	13%	21%	16%
7 1	26%	18%	15%	3%	15%	3%	18%	0%	3%	44%	18%	36%	3%
8 1	31%	5%	3%	8%	3%	28%	15%	5%	3%	36%	10%	51%	3%
9 2	8%	8%	8%	5%	3%	39%	21%	3%	5%	16%	13%	66%	5%
10 2	18%	0%	11%	0%	18%	29%	21%	0%	3%	18%	11%	68%	3%
11 1	21%	21%	5%	5%	28%	5%	15%	0%	0%	41%	10%	49%	0%
12 1	5%	49%	10%	5%	0%	15%	5%	0%	10%	54%	15%	21%	10%
13 2	24%	21%	24%	5%	5%	8%	11%	3%	0%	45%	29%	26%	0%
14 1	3%	0%	0%	0%	5%	44%	23%	0%	26%	3%	0%	72%	26%
15 2	0%	8%	0%	0%	5%	68%	18%	0%	0%	8%	0%	92%	0%
16 1	3%	5%	0%	3%	0%	51%	5%	0%	33%	8%	3%	56%	33%
17 2	0%	3%	3%	5%	0%	47%	5%	3%	34%	3%	8%	55%	34%
18 2	16%	55%	16%	5%	5%	0%	3%	0%	0%	71%	21%	8%	0%
19 2	53%	42%	5%	0%	0%	0%	0%	0%	0%	95%	5%	0%	0%
20 1	85%	0%	3%	13%	0%	0%	0%	0%	0%	85%	15%	0%	0%

* Subscript 1 or 2 following each Icon ID number indicates Booklet number

Icon Appropriateness Rankings

	1	2	3	4	5	6	7
1 1	23%	45%	13%	19%			
2 1	20%	21%	48%	11%			
3 2	45%	24%	29%	1%			
4 2	12%	9%	9%	69%			
5 1	12%	45%	16%	10%	10%	7%	
6 2	21%	15%	38%	12%	10%	4%	
7 1	48%	21%	12%	3%	4%	12%	
8 1	14%	10%	10%	51%	11%	5%	
9 2	1%	8%	11%	11%	47%	22%	
10 2	4%	1%	14%	16%	18%	47%	
11 1	1%	5%	8%	14%	30%	35%	6%
12 1	27%	21%	26%	16%	9%	1%	0%
13 2	4%	0%	5%	6%	6%	17%	61%
14 1	35%	19%	27%	6%	1%	10%	0%
15 2	4%	6%	5%	12%	29%	19%	25%
16 1	18%	39%	21%	12%	8%	3%	0%
17 2	10%	9%	9%	34%	19%	13%	5%

CONCLUSIONS

Based primarily on the data presented in Table 3, the following conclusions are warranted from the icon comprehension testing.

FORWARD COLLISION WARNING SYMBOLS. With reference to Table 3, Icon #3 (shown in Figure 2) achieved the highest comprehension levels (45%) among the four (4) forward collision warning symbols tested, as well as the highest percentage (45%) of #1 rankings. It was also associated with only 5 percent critical confusions.

Figure 2. Icon #3: Symbol with the highest comprehension and ranking scores for the forward collision warning symbols tested.

SIDE COLLISION WARNING SYMBOLS. With reference to Table 3, Icon #6 – closely followed by Icon #7 – achieved the highest comprehension levels (50%, with 44% for Icon #7) among the six (6) side collision warning symbols tested. Icon #6 and Icon #7 are shown below, respectively, in Figures 3 and 4. Icon #6 was associated with 16 percent critical confusions; Icon #7 was only associated with 3 percent critical confusions. Interestingly, Icon #7 had the most #1 rankings among the side collision warning symbols – 48 percent – compared to 21 percent of #1 rankings for Icon #6.

Figure 3. Icon #6: Symbol with the highest comprehension scores for the side collision warning symbols tested.

Figure 4. Icon #7: Symbol with the most #1 ranking scores of the side collision warning symbols tested.

LANE DEPARTURE WARNING SYMBOLS. With reference to Table 3, Icon #12 (shown below in Figure 5) achieved the highest comprehension levels (54%) among the seven (7) lane departure collision warning symbols tested.

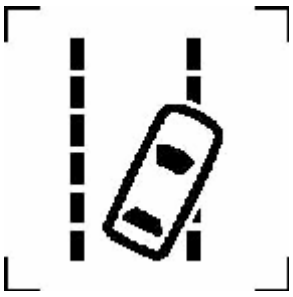

Figure 5. Icon #12: Symbol with the highest comprehension scores of the lane departure collision warning symbols tested.

Icon #11 and Icon #13 had the next highest comprehension scores for the lane departure symbols, with scores of 41 percent and 45 percent, respectively. These icons are shown below in Figures 6 and 7.

Figure 6. Icon #11: Lane departure collision warning symbol with 41 percent comprehension score.

Figure 7. Icon #13: Lane departure collision warning symbol with 45 percent comprehension score.

Interestingly, Icon #14 had the most #1 rankings among the lane departure collision warning symbols – 35% – , but the lowest comprehension (3%) and a high percentage (26%) of critical confusions.

COMPARISON OF TESTING RESULTS ACROSS ISO MEMBER COUNTRIES.

How did the U.S. testing results compare with the results from the other ISO member countries? In brief, results from the other ISO member countries were very similar and comparable to – though not identical to – the U.S. results. Within each of the three classes of in-vehicle active safety system symbols tested, the same (or similar) icons typically placed first or second in both comprehension and rank order testing across all or most of the four studies. Overall, the results from all four studies have provided a solid empirical basis for the selection of active safety symbols within ISO. It is clear from these efforts that the method provides an easy-to-use, yet technically vigorous, set of procedures for international comprehension testing of in-vehicle symbols that can be put to good practice. The current effort, involving numerous ISO member countries, represents an important step forward with respect to the safe and effective development of active safety systems within the international community of vehicle designers.

ACKNOWLEDGMENTS

Battelle's original contract to develop human factors design guidelines for in-vehicle icons was funded by the FHWA (Contract No. DTFH61-97-R-00061); Thomas Granda, Ph.D., was the Contracting Officer's Technical Representative (COTR). The icon comprehension testing reported in this paper was also funded by the FHWA (Contract No. DTFH61-02-C-00134), Task Order BA34006; Ray Resendes of the DOT's Joint Program Office was the COTR.

REFERENCES

1. Balbale, U., & Kiefer, R. J. (June, 2001*). Driver preference rankings of various collision avoidance-related symbols*. Presented at ISO TC22/SC13/WG5 (Road Vehicles-Symbols for Controls, Indicators and Telltales) meeting in Dresden, Germany.
2. Campbell, J. L., Richman, J. B., Carney, C., & Lee, J. D. (2002). *In-vehicle display icons and other information elements. Task F: Final in-vehicle symbol guidelines* (Final Report). Seattle, WA: Battelle Human Factors Transportation Center.
3. Green, P. (1993). Design and evaluation of symbols for automobile controls and displays. In B. Peacock & W. Karwowski (Eds.), *Automotive Ergonomics* (chapter 12, pp 237- 268). London, UK: Taylor and Francis.
4. ISO/DIS 9186. (1988). *Procedures for the development and testing of public information symbols*. Geneva, Switzerland: International Standards Organization.
5. Richman, J. B., Campbell, J. L., & McCallum, M. C. (2002). Effective IVIS comprehension research: Context and response scaling methods [CD-ROM]. *Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society*, 1939-1943.
6. Wolff, J. S., & Wogalter, M. S. (1998). Comprehension of pictorial symbols: Effect of context and test method. *Human Factors, 40*(2), 173-186.