Human Factors Evaluation of New Pictograph-Based Messages for Bilingual Variable Message Signs

Smiley, A, PhD, CCPE, Human Factors North Inc.
Smahel, T. CCPE, Human Factors North Inc.
Erwin, S., P.Eng., Ontario Ministry of Transportation

Paper prepared for presentation
at the Human Factors Session
of the 2014 Conference of the
Transportation Association of Canada
Montréal, Québec

ABSTRACT

In response to requirements of the French Language Services Act of 1989, the Ontario Ministry of Transportation (MTO) was required to implement bilingual messages on a new generation of Variable Message Signs (VMS). The goal of this study, carried out in collaboration with IBI Group, for MTO, was to develop a pictograph-based solution for bilingual VMS, by means of applying a human factors test and analysis procedure.

The first stage of the study involved the development of pictograph concepts using a public consultation process with anglophone and francophone drivers, a jurisdictional scan, and an internal team design charette. In the consultation process, participants were asked to create colour drawings, using symbols and as few words as possible, to describe messages such as traffic congestion, lane blockage, road closure, border delay, and severe winds. The jurisdictional scan explored symbols used by other transportation agencies. Using findings from the public consultation and the jurisdictional scan, the internal team design charette was carried to create design concepts that would be carried forward for comprehension evaluation.

In the second stage of the study, ten message sets related to traffic operations (congestion management, blockages and closures), travel time and safety were tested for comprehension using static images of signs shown (in a roadway context) for four seconds on a desktop computer. Each message set included the currently used English message and at least one pictograph-based bilingual alternative. Following the second stage, design decisions were made. Revised and new messages were tested in the third phase.

In the third stage of the study nine message sets related to traffic operations (congestion management, blockages, and closures) were tested. Each message set included the currently used English message and one pictograph-based alternative. In addition, eight pictograph-based safety messages were tested for comprehension.

The TAC recommended level of comprehension for guide sign messages is 75% (1). Overall, in the final stage of the study, text traffic messages scored 75% or above for 5 of the 7 messages and the pictographic message alternative scored this high for only 2 of the 9 traffic messages. However, even absent any feedback to the participants, there was a clear learning process for these novel graphic messages: by the end of the test, 5 of the 9 pictograph messages had a mean comprehension of 75% or over and 2 messages came very close to this level of comprehension, at 74%. The remaining two messages scored 55 to 58% overall, and improved to 65 to 70% by the end of the test. A statistical comparison showed no significant difference between text and pictograph messages for messages first seen at the end of the test.

INTRODUCTION

The French Language Services Act went into effect in 1989 and required bilingual road signing in designated French Language areas of Ontario. An extension was granted for implementation of bilingual messages on Variable Message Signs (VMS), because they presented special safety challenges warranting further study. In 2004 a study was carried out to examine a pictograph-based solution for bilingual VMS (2). Technological developments since 2004 have removed many constraints on the pictograph messages tested in that study. The goal of the current study was to develop a pictograph-based solution for bilingual VMS that built on the findings from the 2004 study (2) and incorporated the technological possibilities introduced in the new generation of VMS available since 2004. This was achieved by means of applying the human factors test and analysis procedure described below. Both traffic operations and safety messages were tested.

METHOD

Experimental Design and Procedures

The experiment was carried out in three phases, as follows:

Stage 1 = Develop design concepts including public consultation
Stage 2 = Test comprehension of first set of traffic and safety messages
Stage 3 = Test comprehension of second set of traffic and safety messages

Stage 1 involved the development of pictograph concepts using a public consultation process with anglophone and francophone drivers, a jurisdictional scan, and an internal team design charette. The consultation process methodology had previously been used to develop symbols for vehicle controls and displays (3) and is also the first step in the ISO (International Organization for Standardization) procedure for the development of public information symbols (4). Participants were asked to create colour drawings, using symbols and as few words as possible, to describe messages such as traffic congestion, lane blockage, road closure, border delay, and severe winds. A jurisdictional scan was subsequently carried out to explore symbols used by other transportation agencies. Using findings from the public consultation and the jurisdictional scan, an internal team design charette was carried out to create design concepts that would be carried forward for comprehension evaluation.

The Stage 2 and 3 comprehension testing was carried out using proprietary software on a computer workstation. Participants were informed that the messages being tested were new versions of existing messages normally displayed on VMS on Ontario freeways.

As part of a previous VMS study (2) a calibration study had been carried out, comparing driver performance in message recall both on-road (VMS seen as a driver in real-world context) and in the laboratory (VMS seen in a roadway context on a static image on a laptop computer). The goal was to determine the appropriate display interval to use in the laboratory in order to simulate the amount of time available to drivers in the real world to read VMS signs. Based on this research it had been determined that a display time of four seconds in the laboratory was the appropriate interval to attain the same level of recall performance as found in the real world. This display interval was used for computer based testing in the current study.

In Stages 2 and 3, all messages were shown within a roadway context (see Figure 1).

Figure 1: Example test slide: VMS shown in roadway context

In Stage 2 of this study, 25 messages including seven message sets related to traffic, two to travel time and one to safety were tested. There were between one and three bilingual pictograph alternatives for each message set.

In Stage 3 of this study 24 messages, including seven text traffic and nine pictograph-based traffic messages, as well as eight new pictograph-based safety messages were tested.

Participants were assigned to one of six orders. Orders were counterbalanced to ensure that test messages were viewed an equal number of times towards the beginning, middle, and end of the test session. In Stage 2, half of the participants saw messages with lane marks and half without.

Prior to testing, all participants were shown a series of slides illustrating how the collector, express and transfer lanes function. In addition, in order to become familiarized with the test procedure, all participants carried out three practice tasks. During the course of the testing, participants did not receive any feedback as to the correctness of their answers.

Participants

In Stage 1, 36 qualified participants (i.e., valid G license, drive on 400-series highways at least once a week, equally distributed across 3 age groups) were recruited in a mall to provide their input regarding the design of bilingual pictograph messages. All participants had English as a first or second language. In addition, four qualified francophones were recruited for a separate session administered in French.

In Stage 2, 324 participants, evenly distributed between Toronto, Ottawa and Sudbury, were recruited in malls, and then tested on desktop computers in an office setting. For each location one-third of participants were distributed in each of the following age

groups: less than 25, 25 to 55, and over 55 years, with each age group comprising half males and half females. In addition, one-third of participants were in each of the following language groups: English as a first language, French as a first language, neither English nor French as a first language. All participants were required to hold a valid driver's licence (G2 or G) and drive on 400-series highways at least once a month. The test was administered in English for the English language group and French for the French language group. The group whose first language was neither English nor French were able to choose whether they wanted to do the test in English or French.

In Stage 3, a new set of 324 participants were recruited in malls and then tested on desktop computers in an office setting. While testing had been carried out in three cities in Stage 2, the decision was made to reduce this to two for Stage 3, thereby increasing the sample from Toronto. It is Toronto drivers who will be most exposed to the traffic messages tested. Two-thirds of the sample was from Toronto and the remainder was from Ottawa. For each location one-third of participants were distributed to each of three age groups: less than 25, 25 to 55, and over 55 years, with half male and half female in each age group. In addition, one-third of participants were in each of the following language groups: English as a first language, French as a first language, neither English nor French as a first language. All participants were required to hold a valid driver's licence (G2 or G) and drive on 400-series highways at least once a month. Similar to Stage 2, the test was administered in English for the English language group and French for the French language group. The group whose first language was neither English nor French were able to choose whether they wanted to do the test in English or French.

Analysis

Stage 1 was used to develop concepts for testing; performance was not measured.

In Stages 2 and 3 two measures were used to score message performance: units of information retained and message comprehension. Scores for these two measures were determined based on participant responses to a single question with two parts: "(1) What did the sign say and (2) what should drivers do?". Participants were instructed to respond to both parts of the question.

A unit of information answers questions such as: "what is the problem?", "where is the problem located?", or "when will drivers be affected?". For example, the message EXPRESS MOVING SLOWLY, COLLECTOR VERY SLOW contains two units of information. Participant responses were scored as either reporting the information or not (score of 0, 1 or 2 if the message had 2 units of information).

For message comprehension, responses were scored fully correct, partially correct, or incorrect (score 2, 1 or 0). For example, for the message EXPRESS MOVING SLOWLY, COLLECTOR VERY SLOW, a fully correct response would require participants to state that express is preferable to collectors and they should take action to avoid the heaviest congestion (e.g., "stay in express" or "get out of collector"); a response would be considered partially correct if respondents indicate that they are aware of the condition but describe a passive response (e.g., "slow down" or "expect a delay"); an incorrect response demonstrates a lack of understanding of the displayed message (e.g., "use collector" or "do nothing").

A composite score was calculated for each message by equally weighting comprehension (on a scale of 0-2) and units of information retained (on a scale of 0-2).

Note that the number of units of information for messages containing more than two units was pro-rated to a maximum score of 2. For example, if a message had 3 units of information of which the subject reported 2 units, the score, out of a full score of 2, for information transmitted would be 2/3 X 2 = 1.33.

Although the results were presented for comprehension, information transmitted and a composite score including both variables, the discussion below with regard to statistical significance is based mainly on comprehension (percent partially or fully correct) for individual messages. The reason for this is the over-riding importance of comprehension relative to information transmitted, as well as the considerable correlation between comprehension and information transmitted (88%). An ANOVA analysis was required to examine city, language and education effects (not discussed in this report). This can only be carried out using a continuous variable. Thus analyses of the composite score (comprehension and information transmitted) were used for this purpose.

A variety of statistical tests was used to evaluate the results, including One-Way ANOVA, Repeated Measures ANOVA, Chi-Square, and Paired T-Tests.

RESULTS

Stage 1 – Concept Development

Participants were asked to produce design ideas for seven types of traffic and safety messages. The results for each of these are described below.

1. Traffic flow and representation of vehicles

Participants incorporated colours in nearly all of the design concepts. Green was used to illustrate fast, yellow or orange was used to indicate slow, and red was used to indicate very slow or stopped. In some instances arrows were used to illustrate flow, including short arrows for slow speeds and long arrows for fast speeds. Also, boxes were used to illustrate the speed, using colour, or volume (i.e. density) of cars.

2. Travel lanes blocked or closed

Participants used a number of design concepts to illustrate blocked lanes. These included the following:

- Red X alone, rest of lanes empty
- Red X for blocked lanes, green check mark for clear lanes
- Red X for blocked lanes, green up arrow for clear lanes
- Multiple Xs for each blocked lane, rest of lanes empty
- Green arrows to illustrate that drivers must exit the blocked lane

3. Transfer or exit lanes blocked or closed

The following design concepts were used:

- Red X alone, open lane unmarked
- Red X for blocked lane, green check mark for clear lane
- Red X for blocked lanes, green up arrow for clear lane
- Several red boxes to depict stopped cars in blocked lane

4. Full highway closure

The following design concepts were used:

- Red X's for all lanes, green arrow for must exit location
- Red X's for all lanes, green arrow for must exit location and for highway re-entry location

5. Blocked lanes on adjacent highway

The following design concepts were used:

- Adjacent highway drawn parallel to current highway
- Adjacent highway drawn perpendicular to current highway
- Adjacent highway drawn as exit from current highway
- Adjacent highway drawn in box next to current highway

6. Delay at border crossing

The following design concepts were used:

- Clock or stop sign with delay time inside
- Bridge and flags to illustrate border
- Flags, border, and vehicle line-up with delay shown numerically

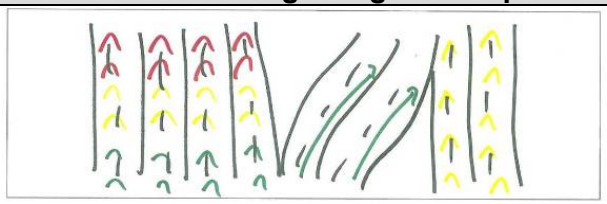
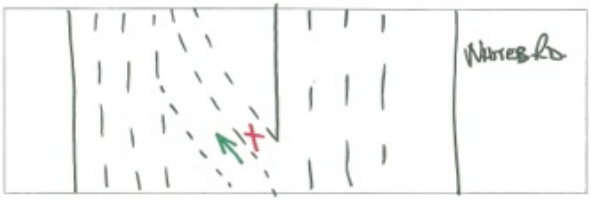
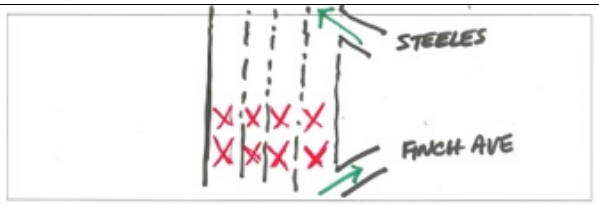
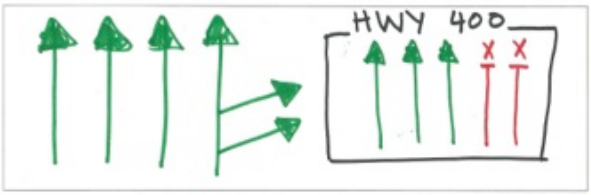
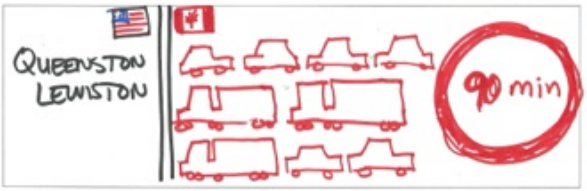
7. Severe Wind

The following design concepts were used:

- Cloud with face blowing on bridge
- Clouds with wind and caution triangle
- Wind sock shown with bridge
- Arrows and wind speed shown in amber

Table 1 illustrates the most promising design concept for each of these seven scenarios.

Table 1: Design Concepts Developed in Phase 1

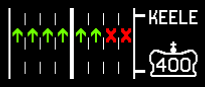
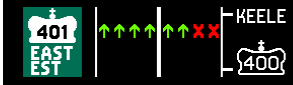
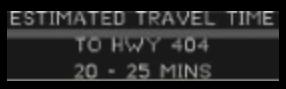
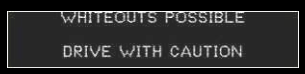
Message Type	Most Promising Design Concept
Traffic Flow	
Lane Blocked	
Transfer Lane Blocked	
All Lanes Blocked – Detour Required	
Blocked Lanes on Adjacent Highway	
Border Delay	
Severe Wind	

Stage 2 – First Comprehension Test

In Stage 2, seven message sets related to traffic, two to travel time and one to safety were tested. Table 2 shows the overall mean comprehension for the current English format and the best bilingual alternative. As noted earlier, participants were assigned to one of six orders. Orders were counterbalanced to ensure that test messages were

viewed an equal number of times towards the beginning, middle, and end of the test session. To show the size of the learning effect, mean percent comprehension is shown for the portion of the sample (one-third) that first saw that message towards the end of the test.

Table 2: ***Stage 2: Summary of comprehension results – English and Best Bilingual Alternative***

Message Type	Current English Format			Best Bilingual Alternative		
	Message	Mean	End of Test	Message	Mean	End of Test
Congestion Management		90%	96%		57%	74%
Full Closure		87%	87%		51%	60%
Partial Blockage		80%	81%		58%	74%
Partial Blockage Adj. Freeway		28%	37%		13%	7%
Lane Transfer Blockage		51%	56%		45%	44%
Use Express Only		77%	81%		49%	56%
Use Collector Only		78%	81%		63%	75%
Travel Time		87%	96%		92%	93%
Border Delay		67%	67%		46%	50%
Whiteout Warning		95%	92%		86%	85%
TOTAL		**74%**	**77%**		**56%**	**62%**

A Chi-square test compared comprehension of the 10 English messages with the 10 best bilingual alternatives. English messages were significantly better understood overall (mean 74% English vs. mean 56% best bilingual alternative, $p < .001$).

Paired T-tests were carried out to determine if the French and Other language groups performed at a higher level with the bilingual pictogram messages as compared to the Current English message. The French group performed at a higher level with the Current English format as compared to the bilingual pictogram alternatives ($T = 7.60$, $p < .001$). Similarly, the Other language group also performed better with the Current English messages as compared to the bilingual pictogram messages ($T = 12.23$, $p < .001$).

The results for each of these message sets follow below.

Congestion Management

For Congestion Management (EXPRESS MOVING SLOWLY, COLLECTOR VERY SLOW), the English message was much better understood than the three bilingual pictograph equivalents (average 90% compared to a range of 53% to 58%). For the English message, 75% reported the first line, “Express moving slowly” and 58%, the second line, “Collector very slow”. In the pictographic version these concepts were approximately equally reported in all three alternatives. The mean comprehension of the best bilingual alternative was 57% overall and 74% for the third of the sample who first saw this message at the end of the test.

Full Closure

With the English message (ALL LANES CLOSED, EVANS TO HWY 427), subjects were much more likely (88% vs. 33%) to report “all lanes closed” than for the bilingual alternatives. Reporting of “from Evans” and “to 427” was more similar for the English message and the alternatives, but still better for the former.

The simpler version of two alternatives, with a single green arrow indicating where to exit, and no arrow indicating where to return, was better understood by participants (47%). There was no effect of the presence or absence of lane marks. When the composite score was examined, the single green arrow version with lane marks performed best.

An examination of the incorrect responses showed that although most participants correctly interpreted the sign as meaning that Evans was accessed on the right, they failed to mention that the road was closed ahead.

The mean comprehension of the best bilingual alternative was 51% and 60% for those who saw this message at the end of the test.

Partial Blockage

The information “COL” meaning “in collector” (COL: 2 RIGHT LANES BLOCKED, BEYOND HWY 400, TO KEELE), despite being the first information read in the English message, was not well retained (36% English message and average of 32% for pictographs). The information “2 right lanes blocked” was retained by 62% of subjects for the English message, but only 42% for the pictograph alternatives. As was the case for full closure, the simpler version of the partial blockage message, without the diversion arrows, and thus with less information to process in a short time, was better understood (53%). A U.S. study (5) found drivers preferred having the diversion arrows included,

indicating that preference does not necessarily guarantee better performance. The mean comprehension of the best bilingual alternative was 58% and 74% for those who saw this message at the end of the test.

Partial Blockage on Adjacent Freeway

The same message of partial blockage was tested, but this time indicating that the blockage was on an adjacent freeway (401 EAST, 2 RIGHT LANES BLOCKED, HWY 400 TO KEELE). A great effort was made to make the alternate freeway information different from the "own" freeway information, by using a very conspicuous crown panel on the right of the sign and telling respondents in advance that they were on Highway 427. Nonetheless, applying the information to the adjacent freeway decreased comprehension considerably from 80% to 28% for the English language message, and from an average of 53% to 13% for the pictograph alternatives. The mean comprehension of the best bilingual alternative was 13% and 7% for those who saw this message at the end of the test.

This message category was the most poorly understood of all, both in its English message form (28%) and in the best pictograph (13%). The typical errors made by subjects were a lack of recognition that the message concerned a freeway other than the one they were on. This could lead to unnecessary lane changing for drivers in the right lanes on Highway 427, and potentially sudden lane changes back into the right lanes for those drivers who were intending to exit shortly to access the adjacent freeway.

Transfer Lane Blockage

At 51% comprehension, the English language version of this message (NEXT COLLECTOR TRANSFER, 1 RIGHT LANE BLOCKED) was the second most poorly understood English message. One reason this may be poorly understood is that the express collector system, involving transfer lanes back and forth, is an unusual highway design. Drivers learn it by driving it, and the word "Transfer" is not used on any permanent signs, only the VMS. Comprehension was poor even though participants were given some instruction on how the collector/express lane system functions, to help them better imagine the context, as they might be expected to do on the actual roadway. The mean comprehension of the best bilingual alternative was 45% and 44% for those who saw this message at the end of the test.

The images shown to subjects during the instruction about the express/collector system were driver's eye view photographs, whereas the image shown in the pictograph was a bird's eye view. Some subjects commented that it would have been easier to understand if the transfer lane that was shown was more diagonal, perhaps because it would have looked more like the driver's eye view.

Use Express Only

The two alternatives to the English only message (EXIT TO 404/DVP, USE EXPRESS ONLY) were a pictographic message and a bilingual message with text. The bilingual alternative using only text was significantly better understood than the pictograph-based (49% vs. 30%). The incorrect answers do not give much guidance as to how to modify the sign. For the English message, the first ("Exit to 404/DVP") and second lines ("Use Express Only") were equally well retained (78% and 75% respectively). The mean comprehension of the best bilingual alternative was 49% and 56% for those who saw this message at the end of the test.

Use Collector Only

As was the case for Use Express Only, the two alternatives to the English only message (KEELE, DUFFERIN, ALLEN, YORKDALE, USE COLLECTOR ONLY) were a pictographic message and a message with text. Again the bilingual text-only version was significantly better understood (63% vs. 40%) than the pictograph-based alternative. The mean comprehension of the best bilingual alternative was 63% and 75% for those who saw this message at the end of the test.

Travel Time

The alternative message tested for travel time used the same words as the English language message (ESTIMATED TRAVEL TIME, TO HWY 404, 20-25 MINS), but fewer of them. As was found in the previous study of bilingual signs, more information presented does not necessarily translate to more information transmitted. Participants understood the simpler message better. There was no effect of using red text to convey the concept of a delay. The mean comprehension of the bilingual alternative was 92% and 93% for those who saw this message at the end of the test.

Border Delay

The English only message for border delay (BORDER DELAY FOR CARS, QUEENSTON 30-40 MINS, FORT ERIE 1-2 HRS) was understood by 67% of participants. The mean comprehension of the bilingual pictograph-based alternative was 46% and 50% for those who saw this message at the end of the test.

Many participants scored as incorrect for the pictograph alternative indicated that the delay was on the highway; they did not mention the border. For the English message 66% indicated that there was a border delay as compared to 28% for the pictographic alternatives. Some subjects suggested an alternate route should be taken.

Whiteout Warning

The alternative message tested for whiteout warning was exceptionally well understood at 86%, not significantly different from the English language message at 95% (WHITEOUT POSSIBLE, DRIVE WITH CAUTION).

Overall Results

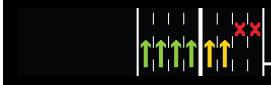
Overall, comprehension reached 75% (or very close) for five of the 10 pictographic messages, for the best alternative for each message. Two of the other five messages are very poorly understood even in the English text version (Partial Blockage Adjacent Freeway (28%) and Lane Transfer Blockage (51%). As for the remaining three messages, Full Closure only reached 60% at the end of the test as compared to 87% for the English text version. Border delay reached 50% at the end of the test as compared to 67% for the English text version. Use Express Only is well understood in the English text version (77%), but only moderately well (56%) in the pictograph message.

Stage 3 – Second Comprehension Test

In Stage 3 seven text traffic and nine pictograph-based traffic messages, as well as eight new pictograph-based safety messages were tested. Table 3 shows the overall mean comprehension for the current English format and the best bilingual alternative. In addition, the mean percent comprehension is shown for the portions of the total sample

that saw that message either at the beginning or at the end of the test. An analysis for each of these message sets follows below.

Table 3: Traffic Messages: Mean Comprehension Compared to the Start and End of Testing

Text Traffic	Start of Test (n=54)	End of Test (n=54)	Mean (n=324)	Pictograph Traffic	Start of Test (n=54)	End of Test (n=54)	Mean (n=324)
EXPRESS MOVING SLOWLY COLLECTOR MOVING WELL	80%	95%	92%		45%	78%	60%
EXPRESS MOVING WELL COLLECTOR VERY SLOW	Not tested				49%	74%	62%
ALL LANES CLOSED BEYOND FINCH	91%	91%	94%		70%	80%	79%
EXP: ALL LANES CLOSED BEYOND MARKHAM	96%	85%	92%		64%	80%	72%
RAMP TO BRONTE RIGHT LANE BLOCKED	69%	62%	64%		83%	74%	74%
EXP: 2 RIGHT LANES BLOCKED COL: 1 RIGHT LANE BLOCKED BEYOND YONGE	80%	76%	74%		38%	70%	55%
HOV LANE BLOCKED 1 RIGHT LANE BLOCKED BEYOND DUNDAS	76%	79%	78%		78%	83%	76%
COLL: 2 RIGHT LANES BLOCKED COLLECTOR MOVING SLOWLY EXPRESS MOVING WELL BEYOND WARDEN	70%	69%	69%		51%	83%	72%
EXP: 2 RIGHT LANES BLOCKED EXPRESS VERY SLOW COLLECTOR MOVING WELL	Not tested				54%	59%	58%
TOTAL	80%	80%	80%	TOTAL	59%	76%	68%

Congestion Management

For Congestion Management, two pictographic versions, one with green and amber arrows, the other with green and red arrows, and one text version equivalent to the former were tested. The English message was much better understood than the bilingual pictograph equivalent (92% compared to 60%). The version with green and red arrows had a similar level of comprehension as the version with green and amber arrows (62% vs. 60%).

Results of very similar versions tested in Stage 2 (red and amber arrows) and Stage 3 (red and green arrows) pictographs were compared and differences were small. As was the case for testing carried out in Stage 2, there was a strong learning effect from the beginning to the end of the session. Average comprehension improved from 80% to 95% for the English language message, from 45% to 78% for the amber and green arrows and from 49% to 74% for the red arrows, overall suggesting that drivers' comprehension would rapidly improve after an initial exposure to the messages. A statistical comparison (based on one third of the sample) showed no significant difference between text and pictograph messages for messages first seen at the end of the test ($n = 755$, $\chi 2 = .183$, $p = .669$, ns).

Some 10% of responses indicated that drivers thought the red arrows indicated a closed road rather than simply heavy traffic. Although this may cause concern it should not. In the real world situation drivers would likely observe and copy other drivers. If large numbers of drivers are continuing ahead and not taking the next off-ramp, a driver uncertain about the meaning of the red arrow is unlikely to conclude that it means that the road is closed.

Participants were asked a number of debriefing questions at the end of the session. Three of these dealt with the meaning of the red, amber and green arrows. Approximately two thirds of the participants were fully correct in their understanding that the colour represented the degree of congestion. Only 1 to 2% had no idea of the meaning and did not even guess. The main concern was the misunderstanding of the red arrows as not being allowed to travel in those lanes (e.g. “closed”, “can’t go”, “danger” or “dead end”). A driver choosing not to return to the collector lanes and detouring by taking the next available exit (which could be two exits before his or her intended exit) would certainly be inconvenienced, but this would not be a dangerous misinterpretation. Furthermore, as noted earlier, drivers are likely to comprehend these messages better when exposed to them in the real world context, where they will be able to see that traffic in these lanes is moving very slowly and there are no other indications, besides the VMS that they have misinterpreted, that the lanes are closed to traffic.

Full Closure

In Stage 2 the Full Closure message (ALL LANES CLOSED, EVANS TO HWY 427) did poorly. A number of changes were made to simplify the message, including removing the exit name following the closure (ALL LANES CLOSED, BEYOND FINCH), using an exit name that was not a highway in an attempt to present a simpler scenario, and angling the depicted off-ramp which allowed for lengthening the arrow indicating the direction drivers were to take. As a result, comprehension for the pictographic message improved, from 46% in Stage 2 up to 79% in Stage 3. In Stage 3, the mean comprehension of the bilingual pictograph alternative was 70% at the start of the test and 80% for those who first saw this message at the end of the test.

Full Closure in Express

The Full Closure in Express message was tested for the first time in Stage 3 (EXP: ALL LANES CLOSED, BEYOND MARKHAM). Comprehension of the text message was 92% as compared to 72% for the bilingual pictograph alternative. Comprehension for the pictograph improved from start to end (a trend) growing from 64% to 80%.

Partial Blockage on Exit Ramp

In Stage 2 Partial Blockage on Exit Ramp was tested for a transfer lane ramp between express and collectors. In Stage 3 a simpler four-lane section version was tested (RAMP TO BRONTE, RIGHT LANE BLOCKED). In Stage 2 comprehension was poor for the English message (51%) as well as for the pictographs (38% to 45%). In Stage 3, with a simpler version of the same message, the pictograph message performed much better than in Stage 2. In addition, the pictograph message actually had significantly ($p<0.001$) higher comprehension (74%) than the English message (64%). For the English message, 10% of the answers were incorrect, specifying that the driver should “take another exit”. Only 1% of respondents indicated this for the pictographic message. This is likely because the pictograph did a better job of conveying that only one of the two ramp lanes was closed.

Concurrent Blockages in Two Traffic Streams

Concurrent Blockage messages were tested for the first time in Stage 3 (EXP: 2 RIGHT LANES BLOCKED, COL: 1 RIGHT LANE BLOCKED, BEYOND YONGE). Comprehension of the text message was 74% as compared to 55% for the pictograph. Comprehension for the pictograph improved significantly ($p<0.001$) from start to end growing from 38% to 70%, and bringing it close to the same level as the text message.

Blockage in HOV and Partial Blockage in General Purpose Lanes

Messages concerning HOV lanes were tested for the first time in Stage 3. The pictograph message was understood almost as well as the text message (HOV LANE BLCKED, 1 RIGHT LANE BLOCKED, BEYOND DUNDAS) (76% vs. 78%). For both messages participants were more likely to recall that the right lane was blocked (incorrect) than that the HOV lane was blocked (correct). This is a little surprising for the text message given that that information was on the first line of the text message. By the end of the test comprehension of the pictograph reached 83%, and of the text, 79%.

In the debriefing questions at the end of the test participants were asked about the meaning of the diamond (restricted lane). About two thirds were fully or partially correct. (There was likely some regional bias in the responses in that diamond-designated lanes are used only by buses in Ottawa.). Given there is no innate connection between a diamond and a restricted lane, it is impressive that participants did so well on this. A total of 60% correctly indicated that the hatched area meant that drivers could not change lanes.

Concurrent Congestion and Blockage

Concurrent Congestion and Blockage messages, with five units of information, were tested for the first time in Stage 3 (COLL 2 RIGHT LANES BLOCKED, COLLECTOR MOVING SLOWLY, EXPRESS MOVING WELL, BEYOND WARDEN). One text message and two pictograph messages (one of which matched the text) were tested. Comprehension of the text version was 69% as compared to 72% for the pictograph alternative (which used amber and green arrows and red Xs). The second pictograph message used green and red arrows and red Xs and it had 58% comprehension.

Both pictographs included red Xs, indicating lane closure. However the version that included a combination of red Xs to mean closure and red arrows to mean very slow traffic may have been confusing. Comprehension of this pictograph did not improve significantly or practically over the test period (54 to 59%), however, that of the other pictograph (amber arrows next to red Xs) did so (51 to 83%, $p<0.0001$).

Answers to the debriefing questions at the end of the test session indicated that 56% of the participants were correct or partially correct in their understanding that the colour of the red arrow represented the degree of congestion. However 24% of participants thought that the red arrows indicated that the lanes were blocked, closed, or inaccessible. With respect to the red X, 86% of participants were correct that these indicated the lanes were closed. This high level was probably in part due to the white arrow indicating the need to leave the closed express lanes by means of the transfer lane.

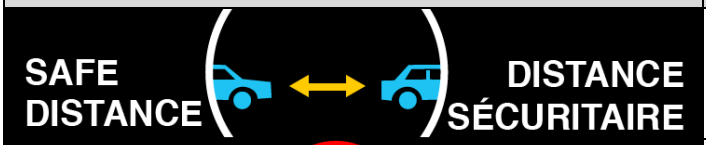
Overall Performance of Pictograph Traffic Messages

The design of the pictograph traffic alternatives was radically different from the current text traffic messages, in that they were pictographs with few if any words, and yet were fully or partially understood by 59% to 83% of participants by the end of the test. Indeed, by the end of the test there was no statistical difference between performance on Text Traffic and Pictograph Traffic messages.

Safety Messages

Table 4 shows the overall mean comprehension for the current bilingual pictograph safety messages tested in Stage 3. Unlike the traffic messages, English text only messages were not tested for comprehension. An analysis for each of these message sets follows below.

Table 4: Safety Messages: Mean Comprehension

Pictograph Safety Message	Mean Comprehension (n=324)
	98%
	96%
	94%
	77%
	32%
	55%
	79%
	83%
TOTAL	**77%**

The SAFE DISTANCE, DON'T FOLLOW TOO CLOSELY, SPEED LIMIT 100 KM/H messages all had 94% comprehension or better.

The SECURE CHILD SEAT message level of comprehension was 77%. The main problem was that participants said "use seat belt" or "secure properly" without mentioning "child seat", and this was scored as incorrect.

The MOVE OVER FOR EMERGENCY VEHICLES message contained three units of information, the critical one of which was "reminder". This scored poorly (32%) because drivers failed to indicate that they did not need to respond immediately. In the debriefing questions at the end of the test, a total of 41% of participants, when presented the options of taking actions now or in the future thought that the message applied now. This misunderstanding is definitely a concern, as it could lead to unnecessary lane changing by a substantial number of drivers, increasing turbulence and the possibility of side swipe crashes.

In the ATTENTION FOG message, the text message did not duplicate the pictograph. The pictograph was selected from a European study as showing best comprehension among a group of alternatives (6). This scored poorly (55%) because drivers failed to mention "fog" or "low visibility". Although the action is correct, the many drivers who said "slow down" or "caution" without mentioning the reason for doing so were scored as incorrect.

The SLOW DOWN, HIGH WINDS ON SKYWAY message contained five units of information. The pictograph contained related but different information from the text (a truck being blown over). Comprehension was 79%.

The BE SAFE, INFLATE TIRES PROPERLY message contained two units of information. The pictograph contained related but different information from the text (tire pressure being measured). Comprehension was 83%.

DISCUSSION

In Stage 1, public consultation, a jurisdictional survey and an internal team design charette provided wide-ranging input and an effective means of developing pictograph concepts for traffic operations and safety messages. In particular, concepts were developed to symbolize traffic flow, lanes blocked, border delays and severe wind.

The TAC recommended level of comprehension for guide or information messages is 75% (1). This was used as a measure of acceptability for the VMS messages.

In Stage 2, with the one exception of the travel time message, English text messages scored higher than pictograph messages. Text traffic messages scored 75% or above for 5 of the 7 message sets. The pictographic message alternative scored this high for none of the 7 traffic messages. However, by the end of the test, 1 of the 7 pictographic messages had a mean comprehension of 75% or over and 2 messages came very close to this level of comprehension, at 74%. No feedback as to the correctness of the responses was given to participants during the test. Thus this learning effect appears to have occurred due to exposure to the full symbol set, e.g. red, amber and green indicating level of congestion. The traffic messages that did most poorly were the text and pictograph messages concerning blockage of a lane on an adjacent freeway (28% and 13%); participants thought the blockage was on their highway. The travel time and

whiteout messages in both text and pictograph versions scored above 75%; neither the text nor the pictograph border delay message scored this high.

In Stage 3, text traffic messages scored 75% or above for 5 of the 7 messages. The pictographic message alternative scored this high for only 2 of the 9 traffic messages. However, by the end of the test, 5 of the 9 pictographic messages had a mean comprehension of 75% or over and 2 messages came very close to this level of comprehension, at 74%. The remaining 2 messages scored 55% and 58% overall, and improved to the range of 65% and 70%, respectively, by the end of the test. The first message showed concurrent blockage events taking place in Express and Collectors. The second message combined red Xs for lane closure with red arrows for traffic moving very slowly.

The red arrows were misunderstood by many participants with 24% thinking that they meant the lanes were blocked, closed, or inaccessible. A greater percentage of participants (86%) did however understand the red Xs as meaning the lanes were closed, blocked, or inaccessible. They were likely assisted by the white arrow indicating the need to use the transfer lane to exit the express lanes. The understanding of the red arrow in the context of traffic moving slowly improved to 74% by the end of the test, suggesting that simple exposure to the system will assist drivers in understanding the intended meaning. Any educational effort should focus on this issue. If it is possible to select visually distinct orange and red colours, consideration could also be given to using orange rather than red for these arrows. It must also be remembered that in a real-life situation drivers will have other cues (e.g. traffic remaining in the lanes) in addition to VMS indicating that lanes are merely moving slowly and not closed (e.g., no police cars, flashing arrowboards).

Of the eight pictograph safety messages, six scored 75% or above for comprehension. Of the two messages that did poorly – Reminder Move Over for Emergency Vehicles, and Attention Fog – only the former is a concern, both because of the very low score and because of inappropriate action (immediate lane changing) that might be taken as a result. This misunderstanding could lead to lane changing by a substantial number of drivers, increasing turbulence and the possibility of side swipe crashes.

CONCLUSION

Public consultation, a jurisdictional survey and an internal team design charette provided wide-ranging input and an effective means of developing pictograph concepts for traffic operations and safety messages. The design of the pictograph traffic alternatives was radically different from the current text traffic messages, in that they were pictographs with few if any words, and yet were fully or partially understood by 59% to 83% of participants by the end of the test. Indeed, by the end of the test there was no statistical difference between performance on Text Traffic and Pictograph Traffic messages. Considering the substantial change in format, and that the participants had never seen such messages before, comprehension was surprisingly good.

ACKNOWLEDGEMENTS

The project was funded by the Ontario Ministry of Transportation. Roger Browne was the contract monitor. Maryann Lovicsek of IBI Group was the project manager for this assignment and Bob Dewar, Conrad Dudek, Milton Harmelink, and Medhat Wahba were key members of the project team.

REFERENCES

(1) Smiley, A. and Dewar, R.E. *Recommended guidelines for testing comprehension*. Report prepared for the Transportation Association of Canada. 2011.

(2) Smiley, A., Donderi, D.C., and Smahel, T. *Bilingual alternatives for VMS, PVMS and static signs*. Final report prepared for IBI Group (for the Ministry of Transportation in Ontario), 2005. Also in the Proceedings of the International Conference on Road Safety and Simulation, Rome, Italy, November 2007.

(3) Green, P. *Development of Pictographic Symbols for Vehicle Controls and Displays*, SAE Paper 790383. 1979.

(4) Zwaga and Easterby. *Developing effective symbols for public information*, in R.S. Easterby and H.J.G Zwaga (Eds.) Information Design (Chichester: Wiley), pp 277-297. 1984.

(5) Ullman, B.R., Trout, N.D., and Dudek, C.L. *Use of graphics and symbols on dynamic message signs: Technical report*. Report No. FHWA/TX-08/0-5256-1. 2009.

(6) Simlinger, P., Egger, S., and Galinski, C. *IN-SAFETY, Infrastructure and safety. Workpackage WP2: Implementation scenarios and concepts towards self explaining roads. Deliverable D 2.3: Proposal on unified pictograms, keywords, bilingual verbal messages and typefaces for VMS in the TERN*. Final report. International Institute for Information Design, IIID, Vienna, Austria. 2008.