Criteria for the Design and Evaluation of Traffic Sign Symbols

ROBERT DEWAR

Several criteria for traffic sign symbols were examined through a questionnaire survey that allowed determination of the importance, or weighting, that should be assigned to each symbol in the design and evaluation of signs. The survey sample included traffic sign experts (members of national traffic control device committees) and practicing traffic engineers from Australia, New Zealand, Canada, and the United States. Separate ratings were assembled for symbols in general and for warning, regulatory, and information symbols in particular. Understandability was the factor rated most important, with conspicuity second. Learnability was considered least important, while reaction time, legibility distance, and glance legibility were rated equally but were determined to be more important than learnability.

The use of symbols (pictographs) to convey information has become prevalent in the past two decades. This is particularly evident in the case of traffic signs, on which symbols are used to convey dozens of different messages. Some of those responsible for traffic control devices believe that almost any message that needs to be conveyed to drivers can be expressed in this form, while others feel that the proliferation of symbolic traffic signs on our highways does more to confuse drivers than to inform them. Recent efforts to develop new symbolic messages indicate that not all messages can be translated into symbols. Research sponsored by the Federal Highway Administration (FHWA) (*1*) indicates that significant proportions of drivers have difficulty understanding symbolic messages that are presently included in the *Manual on Uniform Traffic Control Devices*. In 1985 the FHWA proposed deleting the word message alternates for several traffic signs in the belief that the symbolic versions were well enough understood that the word messages were no longer necessary. It is reasonable to assume that once a symbol has been used on a highway system for many years, drivers will come to know its meaning. This is apparently not the case, however, for many of the symbols presently used.

A task force of the National Committee on Uniform Traffic Control Devices has concluded that certain messages are well understood and do not need to be conveyed with words. The Task Force also believes, however, that evidence on the majority of symbols in the manual is either lacking or indicates that these symbols are not well understood.

Research on traffic sign perception indicates that symbolic messages have a number of advantages over written ones. The most obvious is perhaps the fact that the driver need not be able to read the language of the country in which the symbolic signs are used, which is a benefit for international travelers. Other advantages include greater legibility distance (*2*, *3*), easier recognition under degraded visual conditions such as fog (*4*), readier visibility at a glance (*4*, *5*), possibility of a more rapid response (*4*), and greater conspicuity than word signs (*6*).

It should be noted that these various advantages of symbolic messages reflect several criteria for their effectiveness. Unfortunately, the development of symbolic messages has frequently been hampered by poor research and in some cases no research, as outlined by Dewar and Ells (*7*). Another problem is the tendency to use a single measure of traffic sign adequacy (e.g., understandability, reaction time, or glance legibility) rather than a battery of tests. In some instances, multiple measures have been used (*8*), but even in these studies there has been no indication of the relative importance or weight that should be attributed to each of the measures employed.

In a series of experiments, Dewar and his colleagues used the same set of eight traffic sign symbols and took several measures—legibility distance on the roadway, reaction time, glance legibility, semantic differential ratings, and a preference measure (ratings of clarity of the sign's meaning). Roadway legibility distance was found to be correlated with reaction time (*9*), and semantic differential ratings were correlated with preference ratings (*10*); however, glance legibility was not found to correlate with any of the other measures.

In another series of experiments on traffic signs, Roberts et al. (*11*) used understanding time, accuracy of comprehension, certainty of comprehension, preference, and identification time. An "efficiency index" of each sign's overall effectiveness was calculated on the basis of these five measures. The only meaningful correlation found was that between understanding time and certainty of the accuracy of the response ($r = +0.28$). It appears that the five procedures used by Roberts et al. measured quite different aspects of perception and comprehension of traffic sign symbols.

Another series of experiments, carried out at the University of Melbourne, Australia, also employed several techniques in an extensive evaluation of signs bearing turn restriction messages (*12–15*). Measurements included comprehension, reaction time, glance legibility, legibility distance, and short-term memory for traffic sign messages. Results from the various measures were not always in agreement. Analyses were not performed to determine how the various measures correlated with one another, but they appeared to be measuring different aspects of traffic sign effectiveness.

The various types of research mentioned previously used a number of techniques to measure traffic sign effectiveness. An examination of the results makes it clear that the various

Department of Psychology, University of Calgary, Calgary, Alberta Canada T2N 1N4.

measures are not always closely related. This suggests the need to use more than one method in evaluating traffic sign symbols, but this choice of technique still leaves open the questions of how to combine the data from various measures and what relative importance should be assigned to the different measures. On the basis of previous research on traffic signs, and on the general requirements for a good sign, it is suggested that the following criteria are important in evaluating and designing traffic sign symbols:

- *Legibility distance*. The greatest distance at which the symbol can be clearly "read";
- *Understandability*. The ease with which the symbol can be understood;
- *Conspicuity*. The extent to which a sign can be easily detected or seen in a visually complex environment;
- *Learnability*. The extent to which the meaning of a symbol can be learned and remembered;
- *Glance legibility*. The ease with which the symbol can be "read" when it is seen for only a fraction of a second; and
- *Reaction time*. How quickly the meaning of the sign can be identified.

The study described here examined the relative importance of each of these criteria for the development and evaluation of traffic sign symbols.

METHOD

A questionnaire survey was conducted with eight sample groups of subjects. Four of the groups consist of individuals who can be considered experts in the design and development of traffic signs, and four consist of practicing traffic engineers. The questionnaire asked the subjects to rate, on a 10-point scale from very important to very unimportant, the importance of six criteria for the development and evaluation of traffic sign symbols. Definitions of the criteria, which were listed earlier, were provided on the first page of the questionnaire. The subjects initially rated these criteria without reference to any particular class of traffic sign message. They then rated the same criteria as applied specifically to warning, regulatory, and information signs, assigning separate ratings to each type of sign. Finally, an open-ended question solicited comments on any additional criteria that the subjects might consider important in the design of traffic sign symbols, without reference to sign classification. The questionnaires were distributed by mail to all sample groups except Groups 4 and 8 (described later). For Group 4, the questionnaires were distributed and collected with the assistance of Alan Forbes of the Psychology Department of the University of Wellington (Wellington, New Zealand); in the case of Group 8, the questionnaires were administered at a traffic safety workshop in Sydney, Australia.

SUBJECTS

A total of 153 subjects participated in the survey. All were considered to be knowledgeable about traffic signs and their use on the basis of experience and/or membership on a committee responsible for national traffic sign standards. The sample consisted of four groups of experts and four groups of practicing traffic engineers, as follows:

Group 1. 20 members of the U.S. National Committee on Uniform Traffic Control Devices (NCUTCD);
Group 2. 30 members of the Council on Uniform Traffic Control Devices for Canada (CUTCDC);
Group 3. 11 members of the Standards Association of Australia (SAA) Committee (MS/12), responsible for the Australian Manual on Uniform Traffic Control Devices;
Group 4. 16 New Zealand professionals involved with traffic control devices to varying degrees; five members were on the National Roads Board Committee on Traffic Signs, employees of the Road Transport Division, Ministry of Works and Development of New Zealand;
Group 5. 29 practicing traffic engineers from the United States;
Group 6. 12 practicing traffic engineers from Canada;
Group 7. 21 traffic engineers from Victoria, Australia, who were responsible for traffic control devices in their particular jurisdictions;
Group 8. 14 local government traffic engineers and consultants from various locations in New South Wales, Australia, who were attending a traffic safety workshop.

The sample provides a broad representation of experts and practicing traffic engineers who are highly knowledgeable about the development and design of traffic sign symbols and/or their application to traffic control on the roadways.

RESULTS

The frequency of occurrence of responses to each questionnaire item was determined, and the mean importance ratings were calculated (Tables 1 and 2). Before conducting the major analysis, the reliability of the rating measure and the nationality differences were examined. These preliminary analyses indicated no significant differences between the two Australian samples of practicing traffic engineers, suggesting reliability of the measure, and no differences between the groups of experts from Australia and New Zealand, suggesting that there were no important nationality differences between these two groups. Furthermore, there were no substantial differences between the opinions of practicing traffic engineers and experts. Likewise, differences were minimal between the Canadian and U.S. samples, but for the North American sample the practicing traffic engineers rated four criteria (understandability, glance legibility, and reaction time for warning signs, as well as glance legibility for regulatory signs) as being of greater importance. These small differences indicated good overall consistency in the ratings.

The statistical test used was the median test, which allowed comparison of the particular pairs of samples and pairs of criteria that were of interest. For each analysis the data were divided at the center of the distribution and the chi-square value was calculated. Separate analyses were done for the ratings on traffic signs in general (the first question) and for the individual types of signs—warning, regulatory, and informational. Figure 1 shows the mean ratings of the sample from Australia and New Zealand, as well as those of the Canadian/U.S. sample. To allow comparison of ratings among the six criteria, data from

TABLE 1 IMPORTANCE RATINGS FOR AUSTRALIA/NEW ZEALAND SAMPLE GROUPS

	SAA Committee MS/12	New Zealand Sample	N.S.W. Traffic Engineers	Victoria Traffic Engineers	TOTAL
N	11	16	14	21	62
GENERAL					
Legibility Dist.	3.00	3.69	3.35	3.00	3.26
Understandability	2.27	1.44	1.86	1.76	1.79
Conspicuity	2.64	1.69	1.93	2.14	2.06
Learnability	4.27	3.38	3.86	4.00	3.85
Glance Legibility	3.45	3.31	1.64	2.95	2.84
Reaction Time	2.82	3.06	2.36	2.57	2.69
WARNING					
Legibility Dist.	3.18	3.25	2.36	2.52	2.79
Understandability	2.36	1.75	1.57	1.81	1.84
Conspicuity	2.55	1.69	1.50	2.00	1.90
Learnability	4.18	3.38	3.64	4.05	3.81
Glance Legibility	3.91	3.31	2.14	2.86	3.00
Reaction Time	2.82	3.00	2.07	2.29	2.51
REGULATORY					
Legibility Dist.	2.55	2.63	3.14	3.00	2.85
Understandability	2.27	1.50	2.71	2.00	2.08
Conspicuity	1.82	1.44	2.14	2.10	1.89
Learnability	3.73	2.75	3.57	3.95	3.52
Glance Legibility	2.91	3.00	2.43	3.14	2.90
Reaction Time	2.55	2.38	3.50	2.76	2.79
INFORMATION					
Legibility Dist.	2.91	3.63	2.29	3.95	3.31
Understandability	3.18	2.00	3.36	3.10	2.89
Conspicuity	3.64	2.31	3.36	3.52	3.19
Learnability	4.91	4.50	5.14	4.76	4.81
Glance Legibility	3.55	4.00	3.93	4.81	4.18
Reaction Time	4.27	4.06	4.14	4.62	4.31

* low ratings indicate high degree of importance

all four Australia/New Zealand samples were combined because the preceding analyses had shown essentially the same trends for the four groups of subjects. Within each set of data (general, warning, etc.), all possible combinations of the pairs of criteria were compared. Legibility distance ratings were compared with ratings on understandability, conspicuity, and so on. Similar comparisons were made between the ratings on understandability and the remaining criteria. The same analyses were then carried out for the combined data from the four North American samples.

Results of these analyses (using the median test) are summarized in Table 3, in which only the significant differences are presented. Because of the subjective nature of the measures and the large number of tests carried out, a relatively stringent criterion of $p < 0.002$ was selected as the index of statistical significance for these comparisons.

It is evident that understandability is a particularly important criterion for a traffic sign symbol. Conspicuity ranks a close second behind understandability. Otherwise, the trends were consistent for the Australia/New Zealand sample, although this was not so for the North American sample. The other striking feature is the consistently low rating of learnability. When all the data are considered, the criteria of glance legibility, legibility distance, and reaction time are rated equal to each other in importance, below understandability and conspicuity but above learnability. It should be noted that all criteria are found to be of some importance, if the rating values of 5 and 6 are taken to represent a neutral point on the scale. Three of the criteria approach this neutral rating, however, in the case of information sign symbols.

Some differences can be seen between classes of traffic signs. Understandability appears to be particularly important for warning and regulatory symbols.

The most frequently mentioned additional criteria were sign location (mentioned 18 times), uniformity (18), color (10), night visibility (10), size (6), and shape (6). Note that the first of these (location) is not actually a criterion for sign design but is rather for implementation.

TABLE 2 MEAN IMPORTANCE RATINGS FOR CANADA/UNITED STATES SAMPLE GROUPS

	NCUTCD	CUTCDC	U.S. ENGINEERS	CANADIAN ENGINEERS	TOTAL
N	20	30	29	12	91
GENERAL					
Legibility Dist.	2.95	2.90	3.00	3.25	3.06
Understandability	2.00	2.87	2.07	1.58	2.41
Conspicuity	2.75	2.74	2.86	3.25	2.88
Learnability	3.95	3.97	3.93	3.25	3.90
Glance Legibility	3.15	3.10	2.69	3.00	3.00
Reaction Time	2.60	3.03	2.46	2.08	2.66
WARNING					
Legibility Dist.	2.95	2.60	2.76	3.25	2.82
Understandability	1.84	2.60	2.00	1.25	2.08
Conspicuity	2.70	2.83	2.41	2.58	2.64
Learnability	3.80	3.40	3.62	3.42	3.56
Glance Legibility	3.00	3.70	2.34	2.17	2.92
Reaction Time	2.90	3.03	1.90	2.33	2.54
REGULATORY					
Legibility Dist.	3.20	3.00	3.07	2.92	3.06
Understandability	2.10	2.53	2.00	1.17	2.09
Conspicuity	2.65	2.67	2.62	2.00	2.56
Learnability	3.85	3.30	3.90	2.58	3.52
Glance Legibility	3.25	3.40	2.79	2.92	3.05
Reaction Time	3.00	3.53	3.24	2.50	3.19
INFORMATION					
Legibility Dist.	4.00	2.93	3.79	4.50	3.65
Understandability	2.80	3.16	2.83	2.67	2.91
Conspicuity	3.95	3.57	3.69	4.17	2.67
Learnability	5.32	4.77	4.90	5.17	4.98
Glance Legibility	4.90	3.63	4.10	4.83	4.22
Reaction Time	4.75	4.33	3.72	4.75	4.29

*low ratings indicate high degree of importance

DISCUSSION

The high degree of importance placed on symbol understandability is not surprising. The regularity with which this criterion is incorporated into studies of traffic signs (it is frequently the only variable measured) attests to its importance among researchers. Understandability is dependent not only on how clearly the symbol conveys its intended message but also on the time available for processing it (*2*) and the distance from which it is viewed (*9*). A simple design is recommended because small elements of a symbol cannot be distinguished at the distance usually required in traffic sign perception. In contrast to understandability, the highly rated criterion conspicuity has received very little attention from researchers, except in Australia (*6*). This regional bias may account for the relatively greater importance placed on conspicuity by the researchers in the Australia/New Zealand samples. It could be argued that this criterion is the most fundamental of all, for the other characteristics of a traffic sign symbol become irrelevant if the sign is not seen by the driver. It should be noted that conspicuity per se may not be considered a function of symbol design but is determined more by symbol size, color, shape, and contrast between the symbol and the background of the sign panel on which it appears.

The consistent rating of learnability as less important than the other criteria may be seen by many as a surprise. The low rating of this factor by Group 3, the SAA Committee MS/12 members, is particularly surprising in view of the use at the time of this criterion, along with understandability, to evaluate symbols proposed for the Australian Manual of Uniform Traffic Control Devices. It could be suggested that these results reflect the realization that this criterion is not particularly important in symbol design, especially if a symbol is high in understandability (the most important criterion). In addition, sign designers may feel that learnability is the criterion least

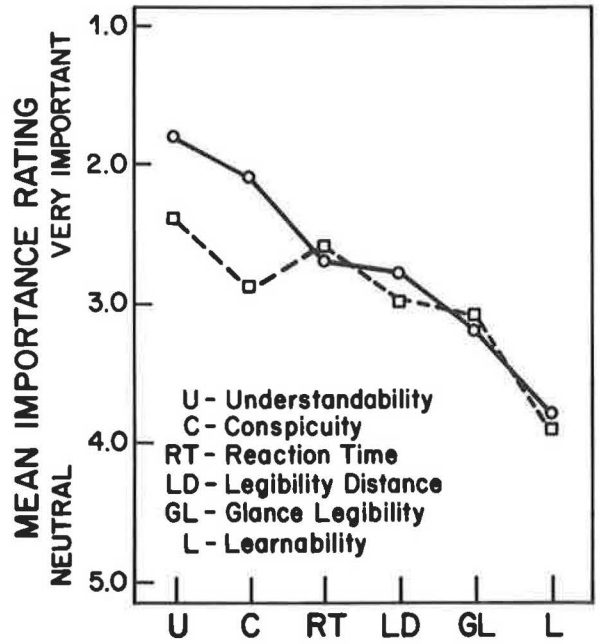

FIGURE 1 Mean importance ratings of six symbol criteria. Dashed line shows ratings from Australia/New Zealand; solid line indicates ratings from Canada/ United States.

TABLE 3 SUMMARY OF SIGNIFICANT DIFFERENCES BETWEEN CRITERIA

Data Set	Australia/New Zealand	Canada/United States
General	U > LD, L, RT C > LD, L	LD > L U > L, GL C > L GL > L RT > L
Warning	LD > L U > LD, L, GL RT > L C > L	U > LD, L, GL RT > L
Regulatory	U > L C > LD, L, GL, RT	U > LD, C, L, GL, RT C > LD, GL
Information	LD > L U > L, GL, RT C > L	U > LD, L, GL, RT C > L

NOTE: U = understandability; C = conspicuity; RT = reaction time; LD = legibility distance; GL = glance legibility; and L = learnability.

under their control, since education of drivers is not their responsibility. However, simplicity of design is often suggested as a worthwhile criterion.

The importance of conspicuity is reflected by the large number of times that sign location or placement is indicated in the spontaneous responses to the open-ended question. Location is not a criterion for symbol adequacy but instead relates to implementation of signing standards by practitioners of traffic engineering. In view of the stress that has been placed on conspicuity, it may be that some subjects see good conspicuity of signs as partial compensation for poor placement. The importance of uniformity of symbols, both within and among traffic sign systems, is also evident. If the stress that has been placed on this issue in the literature is considered, it is surprising that symbol uniformity was not mentioned more often. Visibility under conditions of darkness was of some concern as well. These comments have revealed only one additional criterion (uniformity) that relates directly to design of symbols on traffic signs.

Although committees composed largely of traffic engineers are responsible for determining the designs of symbols for traffic signs, it would be valuable to know the relative importance assigned to symbol criteria by a representative sample of drivers as well. User input has been incorporated into the design of a variety of systems and machines, and the same should be done with visual communication systems used on highways.

The present analysis has shed some light on the issue of the relative importance of the various criteria for traffic sign symbols. The measurement was subjective in nature, and the sample was small and limited to four countries. The overall consistency of the data across the samples, however, permits conclusions to be drawn about the views of traffic sign experts and practicing traffic engineers. It is tempting to suggest the use of a formula with differential weightings applied to each of those criteria, but this would be premature in view of the limited data gathered. However, this study does emphasize the need to take a number of factors into account in the design of symbols. It also provides those who develop traffic signs with information on the relative importance of six criteria for traffic sign symbols.

REFERENCES

1. M. Pietrucha and R. Knoblach. *Motorists' Comprehension of Regulatory, Warning, and Symbol Signs, Volume II.* Technical report. Contract DTFH61-83-C-00136, FHWA, U.S. Department of Transportation, 1985.
2. R. E. Dewar, J. G. Ells, and G. Mundy. Reaction Time as an Index of Traffic Sign Perception. *Human Factors*, Vol. 18, 1976, pp. 381–392.
3. R. E. Walker, R. C. Nicolay, and C. R. Stearns. Comparative Accuracy of Recognizing American and International Road Signs. *Journal of Applied Psychology*, Vol. 49, 1965, pp. 322–325.
4. J. G. Ells and R. E. Dewar. Rapid Comprehension of Verbal and Symbolic Traffic Sign Messages. *Human Factors*, Vol. 21, 1979, pp. 161–168.
5. L. E. King and Z. H. George. A Further Investigation of Symbol Versus Word Highway Signs. Presented at the Annual Meeting of the Human Factors Society, New York, October 1971.
6. B. L. Cole and S. E. Jenkins. Conspicuity of Traffic Control Devices, *Australian Road Research*, Vol. 12, No. 4, 1982, pp. 223–228.
7. R. E. Dewar and J. G. Ells. Methods of Evaluation of Traffic Signs. In *Information Design: The Design and Evaluation of Signs and Printed Material* (R. Easterby and H. Zwaga, eds.). pp. 77–90, John Wiley and Sons, New York, 1984.
8. J. G. Ells, R. E. Dewar, and D. G. Milloy. An Evaluation of Six Configurations of the Railway Crossbuck Sign. *Ergonomics*, Vol. 23, 1980, pp. 359–367.
9. R. E. Dewar and J. G. Ells. A Comparison of Three Methods for Evaluating Traffic Signs. In *Transportation Research Record 503*, TRB, National Research Council, Washington, D.C., 1974, pp. 38–47.
10. R. E. Dewar and J. G. Ells. Methods for Evaluation of Traffic Signs. Final report, Contract 97096. Ministry of Transport, Ottawa, 1974.
11. K. M. Roberts, E. W. Lareau, and D. Welch. *Perceptual Factors and Meanings of Symbolic Information Elements, Volume II.* Report FHWA-RD-77-75, FHWA, U.S. Department of Transportation, 1977.

12. E. R. Hoffmann and W. A. Macdonald. Short-Term Retention of Traffic Turn-Restriction Signs. *Human Factors*, Vol. 22, 1980, pp. 241–251.
13. W. A. Macdonald and E. R. Hoffmann. Information Coding on Turn Restriction Signs. *Proc., 9th Australian Road Research Board Conf.*, Vol. 9, No. 5, 1978, pp. 371–382.
14. W. A. Macdonald and E. R. Hoffmann. Hybrid Turn Restriction Signs: Comprehension of Some Alternative Forms. *Proc., 11th Australian Road Research Board Conf.*, Vol. 11, No. 5, 1982, pp. 105–110.
15. W. A. Macdonald and E. R. Hoffmann. Comprehension of Some Alternative Forms of Regulatory Traffic Signs. *Proc., 11th Australian Road Research Board Conf.*, Vol. 11, No. 5, 1982, pp. 11–26.

Publication of this paper sponsored by Committee on Traffic Control Devices.