

Strength and Understanding of Signal Words by Elementary and Middle School Students

N. Clayton Silver
Department of Psychology
University of Central Florida
Orlando, FL 32816

Michael S. Wogalter
Department of Psychology
Rensselaer Polytechnic Institute
Troy, NY 12180

ABSTRACT

Several recent studies have examined the connoted meaning of signal words that are commonly used in product warning labels and signs. However, the tested population in almost all of these studies has used college students. One purpose of the present research was to determine if the hazard levels implied by signal words connote the same relative meaning to a different population of persons, namely elementary and middle-school students. A second purpose was to assess the understandability of signal words using an objective measure based on the number of missing ratings (i.e., ratings left blank). A third purpose was to develop a list of potential signal words that would be understandable to most younger persons. Elementary and middle-school students rated 43 potential signal words on carefulness (i.e., "How careful would you be after seeing each term?"). A sample of 70 college students also rated the terms on carefulness, strength, and understandability. Although the younger students gave higher carefulness ratings to the words than did the college students, the rank order of the words was consistent across participant groups. In addition, ratings of understandability by college students were predictive of the terms that younger students left blank. Two shorter lists of potential signal words were derived that more than 95% or 99% of the youngest students (fourth and fifth graders) understood. The practical and forensic relevance of these results are discussed, including implications for hazard communication to persons of different populations.

INTRODUCTION

Most standards and guidelines on warning design recommend the use of signal words in signs and labels for the purpose of quickly indicating the level of hazard involved in a particular situation to persons at risk. The standards usually recommend three terms as signal words: DANGER, WARNING, and CAUTION to connote highest to lowest levels of hazard, respectively (e.g., ANSI, 1988; FMC Corporation, 1985). DANGER is intended to connote situations which involve immediate hazards that *will* result in *severe* personal injury or death; WARNING is intended for hazards that *could* result in *severe* personal injury or death; and CAUTION is intended for hazards which *could* result in *minor* personal injury or damage (FMC Corporation, 1985).

However, empirical research on whether people actually interpret different levels of hazard for these words has been equivocal. Some studies have failed to find differences among these and other terms (Leonard, Matthews, & Karnes, 1986; Ursic, 1984; Wogalter, Godfrey, Fontenelle, Desaulniers, Rothstein, & Laughery, 1987). Other research has shown that DANGER implies a greater level of hazard than CAUTION, but has failed to find a difference between WARNING and CAUTION (Bresnahan & Bryk, 1975; Dunlap, Granda, & Kustas, 1986).

In a recent study, Wogalter and Silver (1990) examined 84 potential signal words and made specific comparisons among the three common signal words plus five other terms that had been evaluated in earlier research. They found that the terms DEADLY, DANGER, WARNING, CAUTION, CAREFUL, ATTENTION, NOTICE, and NOTE signified greatest to least strength, respectively. All pairwise differences were significant except between ATTENTION and CAREFUL and between WARNING and CAUTION. The finding of no difference between WARNING and CAUTION supports earlier work (Dunlap et al., 1986; Leonard, Karnes, & Schneider, 1988; Leonard, Hill, & Karnes, 1989), and calls into question the validity of the denoted meanings assigned to signal words as promulgated in guidelines and standards.

However, all but one of the above studies used college students as the population that evaluated the words. The study by Dunlap et al. (1986) used customer engineers and service representatives in the U.S., Canada, and Belgium, as well as U.S. adults without technical training. Significant differences in connotation between signal words were found as a function of population. This result indicates that it is necessary to use groups other than just college students to determine the levels of hazard conveyed by signal words.

In the present study, the population of particular interest was persons of elementary and middle school age. This younger population is at risk of injury because they lack the experience and education of adults. Younger students may not be able to discriminate small differences among the words' connotations. Compared to college students, the younger students' ratings might be higher because all words appear to represent hazard, or they might be in a different rank order. In either case, the reliability of the terms to communicate particular hazard levels would be undermined.

If signal words are to be effective, the population at risk must understand them. Younger persons' lower experience and education makes it more likely that they will not understand less frequently used terms. Thus, it is important for warning research to determine the understandability of words being considered as signal words. Earlier research (Wogalter & Silver, 1990) has evaluated potential signal words using measures such as rated understandability, variability of ratings, and frequency in the language. Another indicant of comprehension is the number of younger persons who are unable to give ratings to terms because they do not understand them. Words left blank by many participants are words that are less likely to be useful in communicating hazard. This missing-values measure is evaluated and compared to the understandability ratings of college students. If the measure is a good one, it could be used to derive a list of words that would be understandable to younger populations.

The purposes of the this research were threefold. The first was to determine whether the terms tested by Wogalter and Silver (1990) are interpreted by elementary and middle-school students in the same way as college students. The second was to evaluate the words using two measures of understandability. The third was to provide a list of terms that virtually all of the younger grade-school students understand for possible use as signal words in warnings.

METHOD

Participants

Two-hundred ninety-eight elementary and middle-school students from Allegheny County, North Carolina (56 fourth, 60 fifth, 50 sixth, 68 seventh, and 64 eighth-grade) participated. Permission from their parents was obtained before they took part in the study. In addition, 70 Appalachian State University (ASU) undergraduates from introductory psychology courses participated in an expanded version of the younger students' questionnaire.

Stimuli and procedure

Forty-three words were selected from a list of 84 terms studied by Wogalter and Silver (1990). Selection from the original list was based on two criteria: (a) word length, and (b) understandability. In the first criterion, words having more than nine letters were excluded on practical grounds because there is limited space on warning labels and signs. A shorter signal word can be made larger than a longer word (which would have a greater likelihood of attracting attention and better visibility at a distance). Also, holding letter size constant, a shorter signal word would provide more space for the other components of a warning (e.g., hazard, consequences, and instructions) than a longer word. In the second criterion, words having mean understandability ratings above 4.0 (rated "understandable" and above on the scale) from the earlier study (Wogalter & Silver, 1990) were included. The purpose of this criterion was to limit the total number of words rated and to have the most understandable words on the list.

Participants were first given the list of terms and told to examine the entire list before starting the ratings. Grade-school students were given a separate sheet that contained the question: "How *careful* would you be after seeing this term?" Below the question was a 9-point scale with the even-numbered anchors having the following verbal labels: (0) not at all careful, (2) somewhat careful, (4) careful, (6) very careful, and (8) extremely careful. Participants were told that the verbal labels were to help them make their ratings and that they could use any whole number from 0 to 8. The instructions emphasized that they should not rate words that they did not understand, that they should leave these words blank.

College subjects rated the terms on the carefulness question and two others on 9-point scales:

- (A) "What is the *strength* of this term?" which had the following numerical and verbal anchors: (0) not at all strong, (2) somewhat strong, (4) strong, (6) very strong, and (8) extremely strong;
- (B) "How *understandable* is this term?" which had the anchors: (0) not at all understandable, (2) somewhat understandable, (4) understandable, (6) very under-

standable, and (8) extremely understandable. For this question, they were to consider whether *all* people in the general population (including young children, visiting foreigners, etc.) would understand the terms.

All participants received one of two random word orders. In addition, the college students each received one of all six possible question orders, and they rated all words on one question before beginning the next question.

RESULTS

Carefulness

Comparison to earlier research. Carefulness ratings were collapsed across subjects for each grade level separately (including the ASU college students) to form mean scores for each word. These scores were then correlated with the mean arousal strength ratings derived from the University of Richmond (UR) students in an earlier study (Wogalter and Silver, 1990). In general, the correlations showed an increase with grade level ($r_s = .78, .86, .93, .90, .94$, and $.94$ for the fourth through eighth graders and ASU students, respectively, $p_s < .0001$). Using a procedure by Rao (1970) for simultaneously evaluating multiple independent correlations, a significant difference among grade levels was noted, $\chi^2(5, N = 43) = 15.52, p < .0005$. A subsequent multiple-range test (Levy, 1976) indicated that the correlation of the fourth grade students' ratings were significantly lower than the correlations of the seventh grade, eighth grade, and ASU college students' ratings ($p_s < .05$). In addition, the ASU students' ratings of strength were strongly correlated with their carefulness ratings, $r = .98, p < .0001$, and the overall arousal strength ratings of the UR students $r = .93, p < .0001$.

Comparison among grade levels and the three common signal words. A 6 (grade level: 4 to 8 plus college students) X 3 (signal word: CAUTION, WARNING, DANGER) analysis of variance (ANOVA) was performed using carefulness as the dependent variable. The means can be found in Tables 1 and 2. The ANOVA showed a significant effect of grade, $F(5, 356) = 2.91, p < .05$. A significant main effect of signal word was also shown, $F(2, 712) = 23.71, p < .0001$. Subsequent Newman-Keuls tests showed that DANGER received significantly higher carefulness ratings than WARNING and CAUTION ($p_s < .01$), with the two latter terms not differing ($p > .05$). There was also a significant interaction, $F(10, 712) = 2.69, p < .01$. Simple effects analysis showed significant differences among the terms for all age groups ($p_s < .05$) except for the fifth and seventh graders ($p_s > .05$). Subsequent Newman-Keuls tests showed that the fourth, sixth, and eighth graders gave significantly higher carefulness ratings to DANGER than WARNING and CAUTION ($p_s < .05$), but the latter two terms did not differ ($p_s > .05$). However, for the college students all three terms differed with DANGER, WARNING, and CAUTION receiving highest to lowest carefulness ratings, respectively ($p_s < .05$).

Comparison among grade levels and the eight previously researched signal words. A 6 (grade levels: 4 to 8 plus college students) X 8 (signal word: NOTE, ATTENTION, NOTICE, CAREFUL, CAUTION, WARNING, DANGER, DEADLY) ANOVA was performed using carefulness as the dependent variable. The ANOVA showed a significant main effect of grade, $F(5, 333) = 4.40, p < .001$. Subsequent

TABLE 1

Mean Carefulness for Words as a Function of Grade Level of Student.

	Grade Level					
	4th	5th	6th	7th	8th	ASU
NOTE	4.80	4.20	2.80	4.00	3.65	3.37
REMINDER	4.96	4.07	3.18	4.09	4.07	3.53
NEEDED	4.94	3.66	3.30	4.37	4.41	4.09
REQUIRED	4.79	4.31	3.85	4.29	4.36	4.27
NECESSARY	5.38	4.26	3.66	4.71	4.36	4.10
ATTENTION	5.16	5.05	3.96	4.64	4.65	4.30
NOTICE	5.25	5.51	3.86	5.12	4.44	4.01
PREVENT	5.84	4.96	4.06	5.33	5.07	4.40
HALT	4.72	5.25	4.22	5.52	4.93	5.61
NO	5.77	5.50	4.20	4.74	4.91	5.60
IMPORTANT	6.31	5.61	4.76	5.25	4.90	5.06
DONT	6.30	5.95	4.26	5.19	5.13	5.24
CAREFUL	5.84	5.88	4.94	5.65	5.20	4.76
PROHIBIT	6.24	5.98	4.89	5.36	5.04	5.30
RISKY	6.65	5.86	5.16	5.46	5.16	5.14
URGENT	6.21	5.06	4.94	5.76	5.55	5.73
ALARM	6.55	5.80	5.38	5.60	5.64	5.01
NEVER	5.71	6.45	4.98	5.79	5.53	5.93
STOP	6.09	6.14	4.80	5.69	5.33	6.43
ALERT	6.75	6.17	5.22	5.84	5.53	5.33
HOT	6.02	5.98	5.12	5.96	5.50	6.21
VITAL	6.29	5.90	5.72	5.80	5.96	5.60
FORBIDDEN	5.94	6.20	5.72	6.06	5.68	5.81
CRUCIAL	6.29	6.46	5.89	5.77	5.90	5.50
INJURIOUS	6.04	6.48	5.98	6.23	5.95	5.37
UNSAFE	6.75	6.20	5.86	6.13	5.64	5.46
CAUTION	6.37	6.90	5.80	6.25	5.88	5.22
BEWARE	6.82	6.50	5.34	6.31	5.97	5.77
SEVERE	6.12	5.85	6.00	6.42	6.45	6.23
WARNING	6.62	6.42	5.86	6.26	5.88	6.13
HAZARD	6.60	6.54	6.10	6.21	6.59	5.67
HARMFUL	6.82	6.67	6.10	6.28	6.48	5.94
SERIOUS	7.17	6.64	6.20	6.45	6.27	5.73
CRITICAL	6.63	6.56	6.53	6.58	6.46	6.03
LETHAL	6.02	6.38	6.27	6.44	6.74	7.41
DANGER	7.38	6.88	6.40	6.57	6.44	6.49
HAZARDOUS	7.00	7.07	6.80	6.94	6.51	6.24
DANGEROUS	7.29	7.08	6.34	6.90	6.52	6.64
FATAL	6.78	7.04	6.63	6.69	7.00	7.36
TOXIC	6.76	6.72	6.88	6.86	7.21	7.17
POISON	7.70	7.30	7.12	7.19	7.12	7.00
EXPLOSIVE	7.54	7.48	7.14	7.54	7.09	7.01
DEADLY	7.89	7.27	7.44	7.65	7.62	7.30
mean	6.26	6.01	5.34	5.86	5.69	5.59
n	56	60	50	68	64	70

Note. Words are ordered according to the overall mean ratings of carefulness (across all participants). Bold and underlined words were analyzed separately.

Newman-Keuls tests showed that: (a) both the fourth and fifth graders gave significantly higher ratings than sixth graders and college students; and (b) the seventh graders gave significantly higher ratings than sixth graders ($ps < .05$). There was also a significant main effect of signal word, $F(7, 2331) = 180.95, p < .0001$, with DEADLY, DANGER, WARNING, CAUTION, CAREFUL, NOTICE, ATTENTION, and NOTE rated from most to least. This order was consistent among all groups, W (Kendall's Coefficient of Concordance) = .97, $p < .01$. Subsequent Newman-Keuls

TABLE 2

Overall Carefulness for Words

	Overall mean	Overall STD	missing ratings
NOTE	3.80	2.57	15
REMINDER	3.97	2.50	12
NEEDED	4.14	2.54	7
REQUIRED	4.31	2.48	22
NECESSARY	4.41	2.61	9
ATTENTION	4.62	2.37	14
NOTICE	4.70	2.41	4
PREVENT	4.93	2.29	14
HALT	5.12	2.35	37
NO	5.13	2.62	3
IMPORTANT	5.30	2.38	6
DONT	5.36	2.48	2
CAREFUL	5.37	2.35	1
PROHIBIT	5.43	2.14	38
RISKY	5.55	2.21	4
URGENT	5.55	2.31	18
ALARM	5.63	2.26	4
NEVER	5.75	2.39	3
STOP	5.78	2.50	1
ALERT	5.79	2.23	6
HOT	5.83	2.42	0
VITAL	5.86	2.27	36
FORBIDDEN	5.90	2.20	19
CRUCIAL	5.92	2.19	48
INJURIOUS	5.98	2.12	41
UNSAFE	5.98	2.18	2
CAUTION	6.05	2.12	6
BEWARE	6.13	2.09	2
SEVERE	6.20	2.25	25
WARNING	6.20	1.97	0
HAZARD	6.27	2.04	7
HARMFUL	6.37	2.02	0
SERIOUS	6.38	1.92	4
CRITICAL	6.44	2.04	28
LETHAL	6.61	2.12	39
DANGER	6.68	1.80	0
HAZARDOUS	6.74	1.87	12
DANGEROUS	6.79	1.70	2
FATAL	6.94	1.82	21
TOXIC	6.95	1.79	13
POISON	7.23	1.67	0
EXPLOSIVE	7.29	1.54	5
DEADLY	7.53	1.30	1

Note. Bold and underlined words were analyzed separately.

tests showed significant differences among all pairwise comparisons ($ps < .0001$) except between WARNING and CAUTION and between ATTENTION and NOTICE. There was also a significant grade X signal word interaction, $F(35, 2331) = 2.01, p < .001$. Simple effects analysis indicated that every grade level showed differences among the words ($ps < .0001$). The means showed the same consistent word order across all grade levels as described above for the main effect. However, there were exceptions: (a) the fifth graders gave CAUTION the second highest mean rating, but the Newman-Keuls test showed that this term was not significantly different from either DANGER or WARNING ($ps > .05$); and (b) the sixth graders, eighth graders, and college students gave NOTICE the second lowest mean

rating, but it was not significantly different from ATTENTION ($ps > .05$). Simple effects analysis also showed significant differences among grade levels for NOTE, NOTICE, CAREFUL, and CAUTION ($ps < .05$), but not for the other words ($ps > .05$). In general, younger students gave higher carefulness ratings than the older students, except that (a) NOTE and NOTICE received significantly lower ratings by the sixth graders than the seventh graders ($ps < .05$); and (b) CAUTION received significantly higher ratings by the fifth graders than the fourth graders ($p < .05$).

Understandability

Correlational analysis. The correlation of the understandability ratings for the college students of both studies was .82 ($p < .0001$) showing good reliability between students of different universities. Of greater importance, however, is whether these ratings reflect actual understanding. An objective measure of the terms' understandability is the (inverse of the) number of missing ratings by the grade-school students. Correlations of the grade-school students' missing data with the ASU students' understandability ratings were -.59, -.64, -.50, -.52, and -.66, for the fourth to the eighth graders respectively ($ps < .0001$). Correlations of the grade-school students' missing data with the UR students' understandability ratings were -.77, -.75, -.62, -.76, and -.74, for the fourth to the eighth graders, respectively ($ps < .0001$). There were no significant linear relationships shown between the understandability ratings of

the ASU students and their carefulness or strength ratings ($ps > .05$).

Comparison of missing ratings among grades. The proportion of missing values for all 43 terms were examined as a function of grade level. College students were not included in the analyses because missing data was extremely infrequent. In general, missing values decreased with increasing grade level ($Ms = .09, .05, .02, .02, \text{ and } .03$, for the fourth to eighth grades, respectively). A one-way ANOVA on these data showed a significant effect of grade level, $F(4, 168) = 34.42, p < .001$. Subsequent Newman-Keuls tests showed that the fourth and fifth grade students had more missing ratings than students of each of the higher grade levels ($ps < .05$). Similarly, the fourth graders had more missing ratings than the fifth graders ($p < .05$).

Lists of understandable terms. Table 3 shows a list of 23 words that the youngest children (fourth and fifth graders) left blank less than 5% of the time. Words denoted by an asterisk were left blank by less than 1% of the youngest children. The mean ratings in this table combine the ratings of the fourth and fifth graders.

DISCUSSION

In general, the order of signal words connoting lower to higher levels of carefulness for the younger students were similar to those of the college students. The pattern of results was also similar to those presented by Wogalter and Silver (1990). One exception was that NOTICE and ATTENTION were reversed in the overall ratings. However, in both studies, these two words did not differ significantly. Another difference was that the ASU students rated WARNING significantly higher on connoted carefulness than CAUTION. While this supports most standards and guidelines (e.g., ANSI, 1988; FMC Corporation, 1985), it does not concur with other research (e.g., Dunlap et al., 1986; Wogalter & Silver, 1990), nor does it correspond with the ratings of the grade-school students in the present research. Why a discrepancy was shown is not clear at this point, and may be due to statistical sampling error.

In general, the results also showed that the younger students gave higher carefulness ratings than the older students. This result might be due a number of reasons. First, the younger students may be less familiar with the words. Although these students might recognize that the words indicate potential hazard, they might not be able to discriminate between different degrees of connotation, and thus rated the words generally higher on the need to be careful. Second, the higher ratings might represent younger students greater regard to possible punishing consequences from protective authority figures (e.g., parents, teachers). That is, not only might these students be worried about the hazardous situation itself, but also with the potential punishment that might result if the warning is not complied with, irrespective of the hazard actually occurring (e.g., spanking, admonishments from authority figures). This additional, and perhaps more likely, consequence of not taking heed to the words, might have instigated the higher ratings. Third, the difference between the younger and older students might reflect greater habituation to the signal words for the older groups (cf. Wogalter & Silver, 1990). Because of greater experience and exposure to the words (and to the situations in which they appear), the words' impact on the

TABLE 3

Carefulness Means and Standard Deviations of Signal Words that were Known by 95% or More of the Fourth and Fifth Graders. Words with Asterisks were Known by 99% or more of the Fourth and Fifth Graders.

Word	Carefulness of 4th and 5th graders	
	Mean	STD
NOTICE	5.39	2.50
NO	5.63	2.78
CAREFUL*	5.86	2.51
IMPORTANT	5.95	2.26
HOT*	6.00	2.59
NEVER	6.09	2.43
STOP*	6.11	2.58
DON'T	6.12	2.30
ALARM	6.16	2.35
RISKY	6.25	2.14
ALERT	6.45	2.02
UNSAFE	6.47	2.18
WARNING*	6.52	2.07
HAZARD	6.57	1.99
CAUTION	6.64	2.03
BEWARE*	6.66	2.04
HARMFUL*	6.74	2.09
SERIOUS	6.90	1.74
DANGER*	7.12	7.12
DANGEROUS	7.18	1.57
POISON*	7.49	1.65
EXPLOSIVE	7.51	1.44
DEADLY*	7.57	1.58

older students might be lower, and thus resulting in lower carefulness ratings. Fourth, the college students might believe they are less likely to get hurt, or at least believe that they are better equipped (e.g., from experience) to handle most hazards than others, therefore to their lower ratings.

The results also showed a strong relationship between the understandability ratings of the college students and the missing ratings of the grade school students. This suggests that both were measuring the same underlying dimension. As might be expected, missing ratings also decreased with increasing grade level. Using the missing values measure, two short lists of words were derived that virtually all of the younger students are likely to understand. These words are some of the most frequently used terms in the English language, and appear to be suitable for warnings that grade-school students and others with lower reading levels might encounter (Johnson, Moe, & Baumann, 1983; Fletcher & Abood, 1988). Fletcher and Abood (1988) noted that over one-half of the signal words on the warning labels are of fourth to tenth grade reading levels. Even so, to minimize potential errors, signal words should be taught to very young children to make sure they know the meaning of the terms (Westaway & Apolloni, 1978).

The current study represents the first time understanding and strength of signal words have been evaluated using younger participants. The reliability that was found between the different populations in this study and earlier research suggests that reasonably accurate predictions can come from initial tests using college student populations. To further establish the reliability of these findings, additional work is underway to examine the connotation of these words with other populations, such as the elderly, recent immigrants, and foreign nationals.

The results have practical and forensic implications. For example, if individuals do not understand the level of hazard implied by a signal word, they may not exhibit the appropriate care and injury may result. If an injury does occur under these circumstances, there may be legal ramifications, including litigation against the manufacturer for not communicating the hazard in a prudent fashion (e.g., *Bituminous Casualty Corp. v. Black & Decker Mfg. Co.*, 1974). A signal word should be chosen to be congruent with the particular hazardous situation. That is, DANGER would not be a wise choice for every potential hazard because although one would be extremely careful, the strength of this word would eventually wear off over time (habituate) because of overuse (cf. Wogalter & Silver, 1990). If a signal word is chosen that minimizes the hazard, then the warning may be inadequate in defining and communicating the potential risk, and the manufacturer would be liable (*Lopez vs. Aro*, 1979). One such example would be using the word NOTE when dealing with harmful, radioactive substances.

Finally, a limitation to the generalizability of the present results should be mentioned. Because the words were presented out of context, it is difficult to generalize to real-world settings. Research on the influence and

connotation of signal words in warnings and in the appropriate environments (e.g., on products and on signs purporting hazardous situations) will assist in determining their utility for communicating various levels of hazard.

REFERENCES

- ANSI (1988). *American national standard on product safety signs: Z535.4-Draft*. New York: Author.
- Bituminous Casualty Corp. v. Black & Decker Mfg. Co. (1974). 518 S.W. 2d 868 (Tex.).
- Bresnahan, T.F., & Bryk, J. (1975, January). The hazard association values of accident-prevention signs. *Professional Safety*, pp. 17-25.
- Dunlap, G.L., Granda, R.E., & Kustas, M.S. (1986). *Observer perceptions of implied hazard: Safety signal words and color words* (Research Report TR 00.3428). Poughkeepsie, NY: IBM.
- Fletcher, D., & Abood, D. (1988). An analysis of the readability of product warning labels: Implications for curriculum development for persons with moderate and severe mental retardation. *Education and Training in Mental Retardation*, 23, 224-227.
- FMC Corporation. (1985). *Product safety sign and label system*. Santa Clara, CA: Author.
- Johnson, D.D., Moe, A.J., & Baumann, J.F. (1983). *The Ginn word book for teachers*. Lexington, MA: Ginn.
- Leonard, S. D., Hill, G. W., & Karnes, E. W. (1989). How does the population interpret warning signals? *Proceedings of the 33rd Annual Meeting of the Human Factors Society* (pp. 550-554). Santa Monica, CA: Human Factors Society.
- Leonard, S.D., Karnes, E.W., & Schneider, T. (1988). Scale values for warning symbols and words. In F. Aghazadeh (Ed.), *Trends in ergonomics/human factors V* (pp. 669-674). Amsterdam: Elsevier.
- Leonard, S.D., Mathews, D., & Karnes, E.W. (1986). How does the population interpret warning signals? *Proceedings of the 30th Annual Meeting of the Human Factors Society* (pp. 116-120). Santa Monica, CA: Human Factors Society.
- Levy, K. L. (1976). A multiple range procedure for independent correlations. *Educational and Psychological Measurement*, 36, 27-31.
- Lopez v. Aro (1979). 584 S.W. 2d. 333 (Tex.)
- Rao, C. R. (1970). *Advanced statistical methods in biometric research*. New York: Hafner.
- Ursic, M. (1984). The impact of safety warnings on perception and memory. *Human Factors*, 26, 677-682.
- Westaway, A., & Apolloni, T. (1978). *Becoming independent: A living skills system*. Bellevue, WA: Edmark Associates.
- Wogalter, M.S., Godfrey, S.S., Fontenelle, G.A., Desaulniers, D.R., Rothstein, P.R., & Laughery, K.R. (1987). Effectiveness of warnings. *Human Factors*, 29, 599-612.
- Wogalter, M. S., & Silver, N. C. (1990). Arousal strength of signal words. *Forensic Reports*, 3, 407-420.