University of North Carolina Highway Safety Research Center

e-archives

Rouse, W.S., and Gendre, F. (1969). Field Experience and Evaluation of TAD Project: Vehicle Damage Rating Scale. Chapel Hill NC: University of North Carolina Highway Safety Research Center.

THE UNIVERSITY OF NORTH CAROLINA
HIGHWAY SAFETY RESEARCH CENTER

FIELD EXPERIENCE AND EVALUATION OF TAD PROJECT "VEHICLE DAMAGE RATING SCALE"

Reprinted from
Proceedings of
Thirteenth Stapp Car Crash Conference

WILLIAM S. ROUSE and FRANCIS GENDRE

DECEMBER 1969

CHAPEL HILL, NORTH CAROLINA

THE UNIVERSITY OF NORTH CAROLINA HIGHWAY SAFETY RESEARCH CENTER
CHAPEL HILL, NORTH CAROLINA

FIELD EXPERIENCE AND EVALUATION OF

TAD PROJECT "VEHICLE DAMAGE RATING SCALE"

Reprinted from
Proceedings of
Thirteenth Stapp Car Crash Conference

William S. Rouse and Francis Gendre

December 1969

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Special recognition and thanks are due Dr. B. J. Campbell, Director of the Highway Safety Research Center, University of North Carolina at Chapel Hill. Dr. Campbell not only conceived the damage rating scale but was also instrumental in introducing the scale into North Carolina. In addition, Dr. Campbell has provided many useful comments and suggestions throughout the entire course of this project. However, any errors in this report are the responsibility of the authors.

Thanks are also due the National Safety Council for furnishing the HSRC with the necessary damage scale manuals. The assistance of the North Carolina Department of Motor Vehicles is also appreciated, and special thanks are due to those members of the Highway Patrol who have been most helpful and cooperative in this project.

REPRINT PERMISSION

The authors wish to thank the Society of Automotive Engineers for permission to reprint this report. The article orginally appeared in Proceedings of the Thirteenth Stapp Car Crash Conference, and a summary of the research was presented at the conference in Boston, Massachusetts in December, 1969.

Reference:

Rouse, W. S. and Gendre, F. Field Experience and Evaluation of TAD Project Vehicle Damage Rating Scale. Proceedings of the Thirteenth Stapp Car Crash Conference, New York: Society of Automotive Engineers, Inc., 1969, pp. 215-236.

ABSTRACT

The TAD vehicle damage rating scale was field tested on a small scale for several months in North Carolina. A test of inter-rater reliability for Highway Patrolmen using the TAD scale showed there is relatively good agreement among raters as to damage type and severity level. Problems in the use of the scale are noted. A second test involved psychological scaling of the TAD manual pictures and showed that the various TAD scales lacked the desired scale characteristics of equal appearing intervals. A sample of 1,329 accident reports were obtained and comparisons made between damage ratings, speed, and estimated damage cost adjusted for vehicle age as predictors of driver injuries. The tests showed that speed is less effective a predictor than either the TAD scale or cost. Cost estimates were as good in predicting driver injury as the TAD scale under some conditions, but the TAD scale proved superior for discriminating serious injuries from minor (or no) injuries.

FIELD EXPERIENCE AND EVALUATION OF TAD PROJECT "VEHICLE DAMAGE RATING SCALE"

The ability to evaluate various vehicle design characteristics and safety devices for crashworthiness on a large scale represents an important goal. Current efforts in this area involve the use of data banks such as a state's accident report file, and the analysis must rely on traditionally reported measures of crash severity (i.e., estimated speed and/or estimated dollar damage to the vehicle) as control variables in equating accident vehicles.

Reported speed and damage cost estimates are thought to lack the precision desired in research of this type and there is some current effort to develop a technique that will provide a better measure of vehicle damage. In addition to providing a good measure of damage, such a tool should also be relatively easy to use and require only a nominal amount of training. It should also require only a few extra minutes of the investigating officer's time and be readily codeable for computer processing.

The National Safety Council has recently published a manual describing a technique for rating damage, titled, "Vehicle Damage Scale for Traffic Accident Investigators." (The manual was produced as part of the Traffic Accident Data project and is frequently referred to as the TAD scale.) This scale seems to meet the requirements listed above and is much less costly than other alternatives such as equipping each patrolman with a camera. Accordingly, this technique of reporting damage is being field tested in a two-county area in North Carolina. This report concerns our experience with the TAD damage rating scale and our evaluation of it.

THE TAD DAMAGE RATING SCALE FORMAT AND TRAINING PROCEDURE

The TAD Project Damage Rating Manual consists of a set of instructions and ten pages of pictures of accident vehicles which represent the more commonly encountered types of crashes (e.g. front-end, sideswipe, roll-over, etc.).* Each page has a set of three pictures showing increasing levels of crash severity represented by increasing amounts of vehicle deformation for a given type of crash. By using the three pictures of a scale as reference points and assigning a number to them as well as points below, between, and above the pictures it is possible to construct a seven-point scale. Ideally, the six lower points on the scale would represent equal appearing intervals. The seventh point is reserved for all amounts of damage above a given level.

One of the principal features of the TAD scale is the identification of not only the location of the damage but also the direction of impact on the vehicle. This feature enables one to determine, to some extent, the post-crash trajectory of vehicle occupants and in turn determine what parts of the vehicle were involved in the "second collision." Thus, for example, damage to the front left corner of a vehicle would be categorized according to whether the impact direction was from the front (which would tend to throw the driver against the steering column), or an angle impact from the side (which would tend to throw the driver in some direction other than against the steering column).

To use the scale one needs only to determine the accident type in terms of damage location and impact direction and turn to the page representing that accident type. The amount of deformation for the crash

---

* The crash categories and code designation are shown in Appendix 1.

vehicle is then checked against the pictorial scale and a crash severity level is determined. The investigator reports the code for the accident type and the number which represents the level of deformation. For example, an "FC-5" represents a front collision with a narrow object such as a tree or utility pole (front concentrated) and indicates a severity level of five on the seven-point scale.

A training session in the use of the TAD scale lasts approximately one hour. Following an introduction to the TAD scale and its purpose, slides showing various types of vehicle damage and damage severity are shown and instructions are given as to what to look for when rating a crash vehicle. After a few slides are shown and rated by the instructor, the class is asked to try rating the rest of the vehicles. This is supplemented by frequent use of prepared drawings and/or blackboard representations of a particular type of collision as the need arises.

In North Carolina, the Highway Patrolmen using the scale were contacted from time to time and a review held of the accidents reported by the individual troopers. During these meetings an attempt was made to correct any misunderstanding in the use of the scale and to find out what kinds of problems arose in using the scale.

INTER-RATER RELIABILITY

In practice, the judgment of the type and level of damage sustained by a vehicle is made on the basis of the investigating officer's opinion alone, and the question arises as to whether or not someone else would judge the damage to be the same. To explore this problem 25 wrecked vehicles were selected and 17 of the original 24 Highway Patrolmen trained to use the TAD scale were requested to rate each of the 25 vehicles.

PROCEDURE - The wrecked vehicles used in this study were selected in the following way.* Two HSRC staff members who were familiar with the TAD scale made independent evaluations of the damage to a vehicle with careful consideration of the TAD damage rating manual. After these independent ratings were made they were compared to determine if there was complete agreement. In case of complete agreement, the vehicle was included in the study. In those cases where there was not complete agreement on the initial rating, a discussion about the damage was held, and if complete agreement could then be obtained the vehicle was included in the study. In two cases (nos. 6 and 16) the damage type was agreed on but the level of damage was rated as "6" by one staff member and "7" by the other one. In these two cases where slight disagreement existed, the vehicles were also included. In those cases where the damage type and severity level could not be easily agreed on, the vehicle was not included in the study. Twenty-two of the 25 vehicles were rated by the HSRC staff members. Vehicles number 23, 24, and 25 were not rated by the HSRC staff. (It should be noted that although the HSRC raters were in agreement this does not mean that they were right.) In addition to selecting wrecked vehicles in which the type and level of damage could be agreed on, an attempt was also made to select vehicles which represented the different types of accidents and various levels of severity. During this selection of vehicles, pictures were taken of the damage for later reference.

---

* The vehicles were selected from among those at the Salvage Disposal Company, Selma, N. C., and thanks are due Mr. Earl C. Helms for allowing this study to be conducted on his property.

The damage ratings were made on two different days by the Highway Patrolmen as their schedule permitted. Those who rated the vehicles on the first day rated all 25. On the second day, two of the 25 vehicles (nos. 2 and 24) had been removed by the salvage company and were not rated by the remaining 7 patrolmen.

The patrolmen were instructed that they would need the TAD manual for the ratings; however, no check was made to determine whether or not they actually used the manual. They were also instructed to make their judgments independently and not confer with others about the damage. Since the patrolmen did not have information as to the circumstances of the accident such as they have in normal accident investigation, they were instructed to do the best they could in determining the accident type in regard to direction of impact. (It should be noted that without knowledge of the circumstances of the accident in some cases it is difficult to determine whether the collision was, for example, a front left or left front quarter.) To help in deciding the direction of impact, the patrolmen were instructed to note the direction in which the metal was bent and any other clues that might be present. The final instruction was that most of the vehicles could be rated for damage using only one scale but a few might require two, or perhaps three, ratings for a complete description. More than three ratings for a vehicle were not permitted.

RESULTS AND DISCUSSION - The ratings first were corrected for obvious errors in which the wrong side of the vehicle was indicated as damaged. This was done by checking the photographs of the vehicles when a discrepancy appeared in the reports. A total of eight errors of this type appeared

with four of them attributed to one individual. A second treatment of data involved eliminating damage reports in which only one or two raters reported secondary or tertiary damage for a vehicle. There were 13 ratings of this type eliminated and all but one of them was a severity level "1." Six of these ratings were reported by one patrolman.

Table 1 shows the data on inter-rater reliability in summary form. Columns 2 through 11 show a summary of the judgments of the majority opinion on damage type for each vehicle, and columns 12 through 14 show the same type of information in reduced form for the minority opinion on damage type. For example, vehicle number 3 (see col. 1) was judged to have two different types of damage as indicated by entries in two rows. The first damage was due to an angle collision at the right back quarter, or an RBQ (see col. 2) according to the judgment of 17 of the 18 raters (shown in col. 10). The remaining raters (in this case "1" as indicated by the "1" in col. 13) judged the damage to be due to a sideswipe-type collision as indicated by an RD entry in column 12. The judged severity of the RBQ damage according to frequency of reports is shown in columns 3 through 9. The severity level most frequently reported for this damage was level "6" (8 respondents), and this modal value plus and minus one severity level was reported by a total of 15 of the 17 raters who reported RBQ damage (see col. 11). The modal severity level (or highest level when no modal value exists) of the damage for the minority opinion is shown in column 12. Each severity level reported by the minority opinion is checked against the mode plus and minus one as reported by the majority for agreement on severity level even though there is disagreement on damage type. In this case the single minority report was severity

TABLE 1

SUMMARY DATA ON INTER-RATER RELIABILITY

1	2	3	4	5	6	7	8	9	10	11	12	13	14
						COLUMN							
	I					MAJORITY				I		MINORITY	I
VEH.	DAM-				SEVERITY				REPORTS	MODE	DAM-		MODE
NO.	AGE	1	2	3	4	5	6	7	RATERS	± 1	AGE	N	± 1
1	FD	0	0	0	6	9	2*	0	17/18	17/17	FL-6	1	1/1
2	FL	0	0	4	4	2*	0	0	10/11	10/10	FD-4	1	1/1
3	RBQ	0	0	1	1	1	8	6*	17/18	15/17	RD-7	1	1/1
	LD	0	0	1	0	3	6*	5	15/18	14/15	LP-5	3	2/3
4	LP	0	4	7	4*	2	1	0	18/18	15/18	-	0	-
5	FD	0	0	3	6*	5	3	1	18/18	14/18	-	0	-
6	LP	0	0	0	0	0	11*	7*	17/19	17/17	LD-7	2	1/2
7	FR	0	2	7	4	0	1	0	14/18	13/14	RFQ-3*	4	4/4
8	FD	0	2	9*	7	0	0	0	18/18	18/18	-	0	-
9	FC	0	2	4*	10	0	0	0	16/18	14/16	FD-3	2	2/2
10	BR	0	0	0	5	6*	2	0	13/18	13/13	BD-4	5	3/5
11	L&T	0	3	9*	4	1	0	0	17/17	16/17	-	0	-
	R&T	3*	2	0	0	1	0	0	6/17	5/6	MIX-1	2(9)	2/2
12	BD	0	0	1	2	6*	1	1	11/18	9/11	BR-6	7	4/7
13	BL	18*	0	0	0	0	0	0	18/18	18/18	-	0	-
14	FR	1	7	5	1	0	0	0	14/18	13/14	MIX-3*	4	4/4
15	BD	0	0	9	7*	0	0	0	16/18	16/16	MIX-4	2	2/2
16	FR	0	0	0	3	0	9*	7*	19/19	16/19	-	0	-
17	LFQ	0	4	6*	1	0	0	0	11/18	11/11	MIX-2	7	5/7
18	LD	0	0	3	6	6*	1	0	16/18	15/16	LP-2	2	2/2
19	FC	0	0	0	5	8*	5	0	18/18	18/18	-	0	-
20	LP	14*	0	0	0	0	0	0	14/18	14/14	MIX-1	4	3/4
21	R&T	0	0	0	3	6*	6	1	16/18	15/16	MIX-5	2	2/2
	LFQ	0	6	2	0	0	0	0	8/18	8/8	MIX-5	4(6)*	2/4
22	L&T	0	1	5*	1	3	0	0	10/18	7/10	MIX-3	7(1)	6/7
	R&T	1	2*	3	0	1	0	0	7/18	5/7	MIX-2	9(2)	5/9
23	RBQ	0	0	0	1	0	2	8	11/17	10/11	MIX-7	6	3/6
24	LFQ	0	0	0	1	0	7	6	14/17	13/14	MIX-5	3	1/3
	LBQ	0	3	5	1	1	0	0	10/17	9/10	MIX-4	3(4)	2/3
	FR	0	5	2	0	0	0	0	7/17	7/7	RFQ-3	3(7)	3/3
25	FR	0	0	2	4	2	1	0	9/10	8/9	FD-3	1	1/1
	FL	0	0	1	3	2	1	1	8/10	6/8	MIX-5	2	2/2
	RD	0	0	0	1	2	1	0	4/10	4/4	RP-5	4(2)	3/4
	PRIMARY DAMAGE								372/434	345/372		61(1)	46/61
								%	85.71	92.74			75.41
	SECONDARY DAMAGE								54/98	47/54		23(21)	15/23
								%	55.10	87.04			65.22
	TERTIARY DAMAGE								11/27	11/11		7(9)	6/7
								%	40.74	100.00			85.71
	TOTAL								437/559	403/437		91(31)	67/91
								%	78.18	92.27			72.72

* HSRC STAFF RATING

"7" and this fell within the severity limits of "5," "6," and "7" reported by the majority. This is indicated in column 14 as one out of one raters agreeing with the majority on severity level.

The second type of damage for vehicle 3, as reported by the majority (15 out of 18 raters) was an LD and 14 of these agreed on the severity level (mode plus and minus one). The minority group (N = 3) judged the damage to be an LP with modal severity of "5" and 2 out of three of these raters judged the severity level at the same level as most of the majority group.

In several cases a particular area of damage was judged to have been due to more than one possible collision type according to independent rater reports. In these cases a "mix" entry is shown indicating mixed opinion.

All of the raters reported some type of primary damage* for each vehicle and for some of the vehicles several raters reported a second or third type of damage. In those cases where some of the judges did not report the secondary or tertiary damage a second entry (enclosed in parentheses) is found in column 13 showing the number of raters who did not report damage.

The data in Table 1 are summarized at the bottom of the Table. It will be noted that there was a total possible number of judgments of 559 (col. 10) and that 437 of the judgments were considered as majority opinion and indicates that about 78% of the time raters were in agreement as to the type of damage. This total figure may be broken down according

---

* Primary damage is defined as the most severe damage according to a majority of the raters.

to agreement on primary, secondary, and tertiary damage as shown in column 10. In the case of the primary damage type, the percentage of agreement jumps to almost 86% but is considerably less for the secondary and tertiary damage. However, when the raters agree on the damage type, it can be seen in column 11 that they are in fairly good agreement as to the level of severity when the range of acceptable severity levels includes the modal value plus and minus one severity level. In the case of total judgments for the majority opinion, the agreement on severity level is 92.27% and about the same for the primary damage. The data at the bottom of column 14 show that when a report of damage type is in conflict with what most others would report there is still a good chance that the reported severity level of the damage would be consistent with the severity level reported by others. In the case of total reports, in column 14 it can be seen that approximately 73% of the time the severity level would be in agreement with the severity level reported by a majority of raters.

These data indicate several things. First, it should be repeated that the patrolmen were required to rate the damage to a vehicle under somewhat artificial conditions since they had no knowledge as to what actually happened in the crash. Even with this handicap they appeared to be in fairly good agreement as to the type of accident (78.18%), and when they agreed on accident type they were even in better agreement as to the severity level (92.27%). In those cases where a rating was not in agreement with the majority opinion as to accident type, the disagreement was due in most cases to a difference in interpretation of the direction of impact on the vehicle. For example, for vehicle number one the majority considered the damage as resulting from a distributed impact on the front

of the car while the lone dissenter judged the impact to be more on the left front. The exact reverse occurs for vehicle number two. Another interesting example is vehicle number 14. In this case the HSRC staff was among the minority in opinion. The actual damage to this vehicle occurred on the front right, as indicated by the majority opinion; however, it was a concentrated impact on this corner and was reported as such by one other rater. The two other raters reported this damage as front distributed. It seems that in general the judgments of the Highway Patrolmen were in good agreement considering the artificial situation created and that a more powerful test of inter-rater reliability would require damage ratings at the accident scene where the circumstances of the crash could be called on in making damage estimates.

The major problem that is evident from the data in Table 1 is that in several cases the judged severity level for a given type of collision ranged over five points of the seven-point scale. This is considered to be undesirable and indicates that the TAD scale needs improvements. Actual experience with the scale and conversations with the Highway Patrolmen indicate that the improvements could come from two sources. First, a better selection of pictures seems warranted, and second, an improved training session with more emphasis on guidelines to be used in rating severity level appears to be needed.

Although it appears that there is fairly good agreement among the raters according to the above analysis, there is still the question as to whether or not certain individual raters systematically deviate from group norms. To study this problem the data were sorted according to raters, and these figures are presented in Table 2.

TABLE 2

INTER-RATER RELIABILITY DATA ON RATER AGREEMENT WITH MAJORITY ON DAMAGE TYPE AND TENDENCY TO OVER OR UNDER RATE DAMAGE SEVERITY LEVEL IN REFERENCE TO MODAL SEVERITY LEVEL.

RATER NUMBER	VEHICLES RATED	AGREES WITH MAJORITY ON PRIMARY DAMAGE	SIGN. LEVEL	OVER/UNDER MODAL SEVERITY RATING	SIGN. LEVEL
1*	22	20	.001	6/1	N.S.
2	23	20	.001	6/2	N.S.
3	23	18	.005	4/3	N.S.
4	23	21	.001	2/2	N.S.
5	25	15	N.S.	8/0	.01
6	25	21	.001	4/3	N.S.
7	23	20	.001	10/1	.02
8	23	22	.001	12/0	.001
9	25	22	.001	1/13	.01
10	25	24	.001	11/1	.01
11	25	23	.001	10/4	N.S.
12	23	20	.001	1/6	N.S.
13	25	20	.002	1/12	.01
14	25	23	.001	5/3	N.S.
15	23	21	.001	2/1	N.S.
16	25	19	.007	3/5	N.S.
17	24	22	.001	4/4	N.S.
18	25	19	.007	3/2	N.S.

*HSRC staff rating

The Table shows the number of vehicles rated by each trooper and the number of times he agreed with the majority opinion as to the type of damage. This test was conducted in order to ascertain whether or not any of the individual troopers were systematically judging the damage type differently from the group as a whole. A one-tailed "sign test" (1)* showed that all of the troopers except trooper number five were significantly different from chance in the assignment of the primary damage type and indicates that they were using essentially the same criteria. Rater number five, who did not seem to be using the same set of criteria in determining damage type, was trained in the use of the TAD damage scale at the same time as the other troopers. However, he has been on special assignment for a number of years and only rarely investigates accidents. During the time since the TAD scale has been in use in North Carolina (about eight months) this trooper has reported only one accident. This indicates that perhaps either some amount of practice may be necessary before a person becomes proficient in the use of the scale, or it is necessary to use the scale fairly frequently in order to stay proficient. However, this interpretation must be taken with caution since other factors may be operating.

A second test of individual performance in relation to the group is also shown in Table 2. The question posed here was to what extent is there a tendency, if any, for some of the raters to rate consistently the severity of damage higher or lower than the group. A two-tailed "sign test" (1) shows that five of the troopers tended to rate the damage as higher or lower than the group modal value at the .01, or .02 level of

---

* Numbers in parentheses designate References at end of paper.

significance and that one rater (no. 8) tended to overestimate the amount of damage to an extent that is significant at the .001 level. The nature of the problem of over-and underestimating damage level has not been explored and the cause is unknown. One possibility, though, is that these men are not carefully consulting the damage rating scale when rating a vehicle. However, whatever the nature of this problem, it is one that should be investigated further and steps taken to correct it.

A final comment must be made regarding the extent of agreement among judges found in this study. The procedure used in selecting vehicles involved a requirement that the HSRC staff raters had to agree on the damage in order for the vehicle to be used in the study. This requirement precluded the selection of a random group of vehicles and may have resulted in selection of a group of vehicles whose damage was relatively unambiguous.

PSYCHOLOGICAL SCALING OF THE "TAD" MANUAL PICTURES

The pictures showing damaged vehicles used in the damage rating manual were selected from a limited sample on the basis of the judgments of two people. It was therefore considered desirable to obtain additional information regarding the apparent amount of deformation for the vehicles in each picture and to have this information for each picture in relation to each of the other pictures. Of principal concern was the extent to which each scale has the characteristics of an equal appearing interval scale.

PROCEDURE - Extra copies of the damage rating manual were obtained from the National Safety Council and the pictures cut out (removing all identifying codes in the process) and stapled to cards.

Each completed set of pictures was randomly arranged and each picture assigned a number for later identification and computer processing.

The subjects were 25 Highway Patrol trainees and thus represented the type of men who would be using the manual. However, the men had not had training in accident investigation at the time of testing.

The subjects were available for one 50-minute class period and were all tested at one time. After a brief introduction to the general problem of evaluating highway safety programs, the men were instructed to sort the 30 pictures into 10 piles according to increasing amounts of damage (deformation) to the vehicle. A work sheet with a ladder-type arrangement of boxes was provided each subject to assist in keeping the pictures in order. The subjects were instructed that there was at least one picture in the set that should be rated at the ten level (most severe) and at least one picture which showed a damage level of "1" (least severe). This instruction was included in order to anchor the two end points of the scale. They were also instructed several times during the testing to reevaluate each picture several times in case some of the pictures needed to be assigned a different rank in view of the assignment of ranks to the other pictures. This part of the testing required approximately 25 minutes.

Following the sort into ten piles, the subjects were instructed to place the number of the pile, in regard to its location on the ten-point scale, on each of the cards in that pile. Following this, the subjects were instructed to take each pile, one at a time, and rank the pictures in that pile in terms of the severity of damage and assign a "1" to the vehicle-picture in the pile with the least damage, a "2" to the next least damage, and so on for each picture in the pile and for each pile. This part of the

testing required approximately 15 minutes. This procedure resulted in a complete ordering of the thirty pictures in two steps.

RESULTS AND DISCUSSION - The range of assigned ranks for each of the pictures grouped according to the TAD manual format are shown in Figure 1. In many cases an overlap exists in assigned ranks for pictures representing a particular scale, and the individual rankings resulted in a high degree of variability for some of the pictures in regard to their relative location on the 30-point ranking scale. For example, the L&T/R&T ranks two and four pictures were judged to be about similar, and the L&T/R&T level-four picture had a spread of seventeen rank points. In only one case (LFQ/RFQ) is there a clear separation between the three levels of severity as shown in the pictures. However, even in this case it seems apparent that the scale is not an equally appearing interval scale because the center picture (damage-level four) is displaced upward rather than being in the center and equal distance from damage levels two and six. The general interpretation of this is that a lot of ambiguity exists in the pictures in terms of how well they portray damage. This may be due in part to highlights and shadows that tend to obscure the damage.

The data were further subjected to a test for statistical significance of the agreement among raters as to relative amounts of damage portrayed in each picture. The test for significance was achieved in three steps. The data were first subjected to a rank correlation procedure which yielded a Mean Rho of .889. This value was transformed into a Kendall's W of .903 (coefficient of concordance) which was in turn transformed into a Chi Square of 654.60 with 29 degrees of freedom (1). A Chi Square value this high is significant well beyond the .001 level and

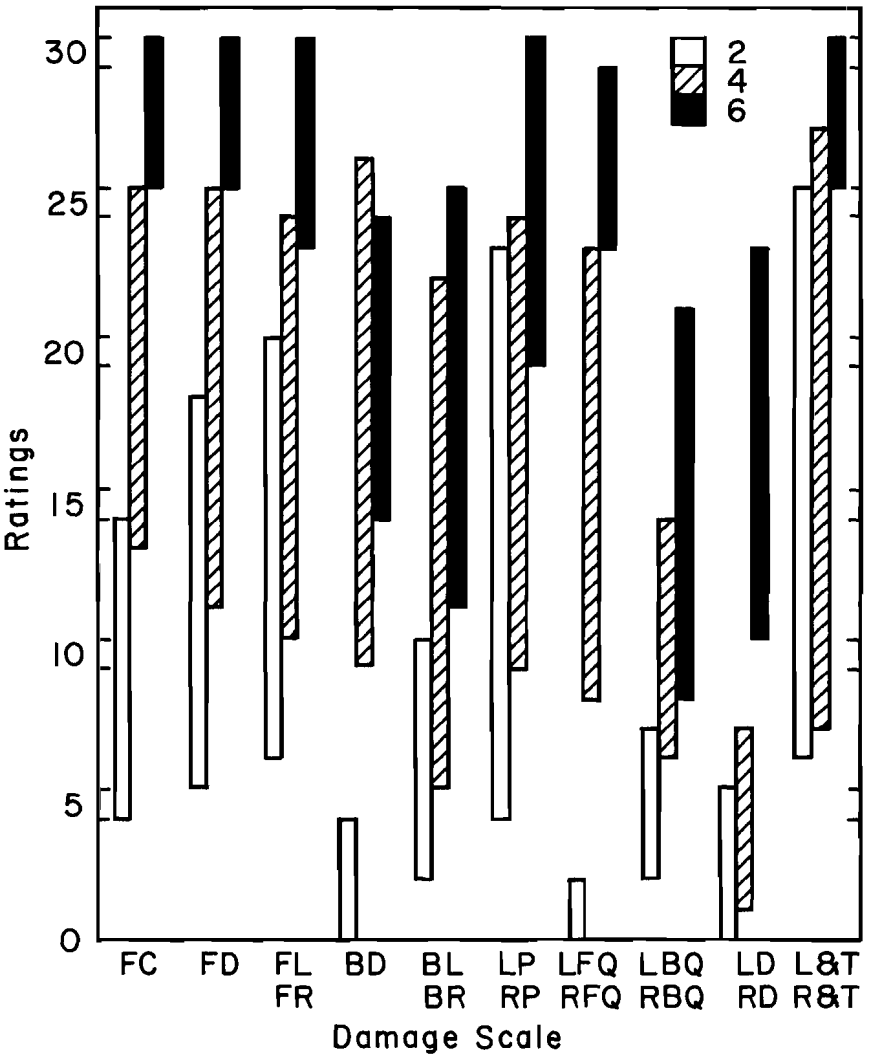

Figure 1. Assigned ranks for each picture of the TAD scale according to results of a psychological scaling procedure.

indicates that the subjects were in considerable agreement as to the relative rankings of the individual pictures on the 30-point scale.

Since the agreement among subjects regarding the relative ranking of the various TAD pictures was good, it was considered desirable to obtain an adjusted rating of the various damage-type severity-level scale intervals for further experimental testing. These adjusted rating scales were obtained by deriving "paired comparisons" (2) and applying the procedure devised by Thurstone (3). The resulting weights are shown in Table 3 and graphically in Figure 2. The values shown in Table 3 have been adjusted by adding a constant to each one in order to avoid negative weights.

A difference of .35 between any two scale values is significant at the .05 level (2). Using this value (.35) the various damage-type severity-level points shown in Part A of Figure 2 may be regrouped into a ten-point scale as shown in Part B of Figure 2. These reassigned values are shown in Table 4 and indicate how a reported damage may be transformed into an adjusted value on a unique 10 point scale. For example, a reported damage of "FC-7" equals a "10" on the revised scale. (References to damage type are not retained in the ten-point scale.)

These reassigned values, obtained on the basis of psychological scaling of the TAD pictures, will be used in experimental testing in the next section of this report and will be designated by the reference "NSc" (new scale).

TABLE 3

DERIVED WEIGHTS FOR EACH OF
THE TAD MANUAL PICTURES

DAMAGE TYPE	DAMAGE LEVEL 2	4	6
FC	1.353	2.856	4.116
FD	1.513	2.875	3.858
FL/FR	1.951	2.722	3.822
BD	0.000	2.424	2.747
BL/BR	0.716	1.802	2.830
LP/RP	1.238	2.186	3.692
LFQ/RFQ	0.109	2.297	3.673
LBQ/RBQ	0.427	1.473	1.949
LD/RD	0.398	0.164	2.253
LT/RT	2.314	2.649	4.087

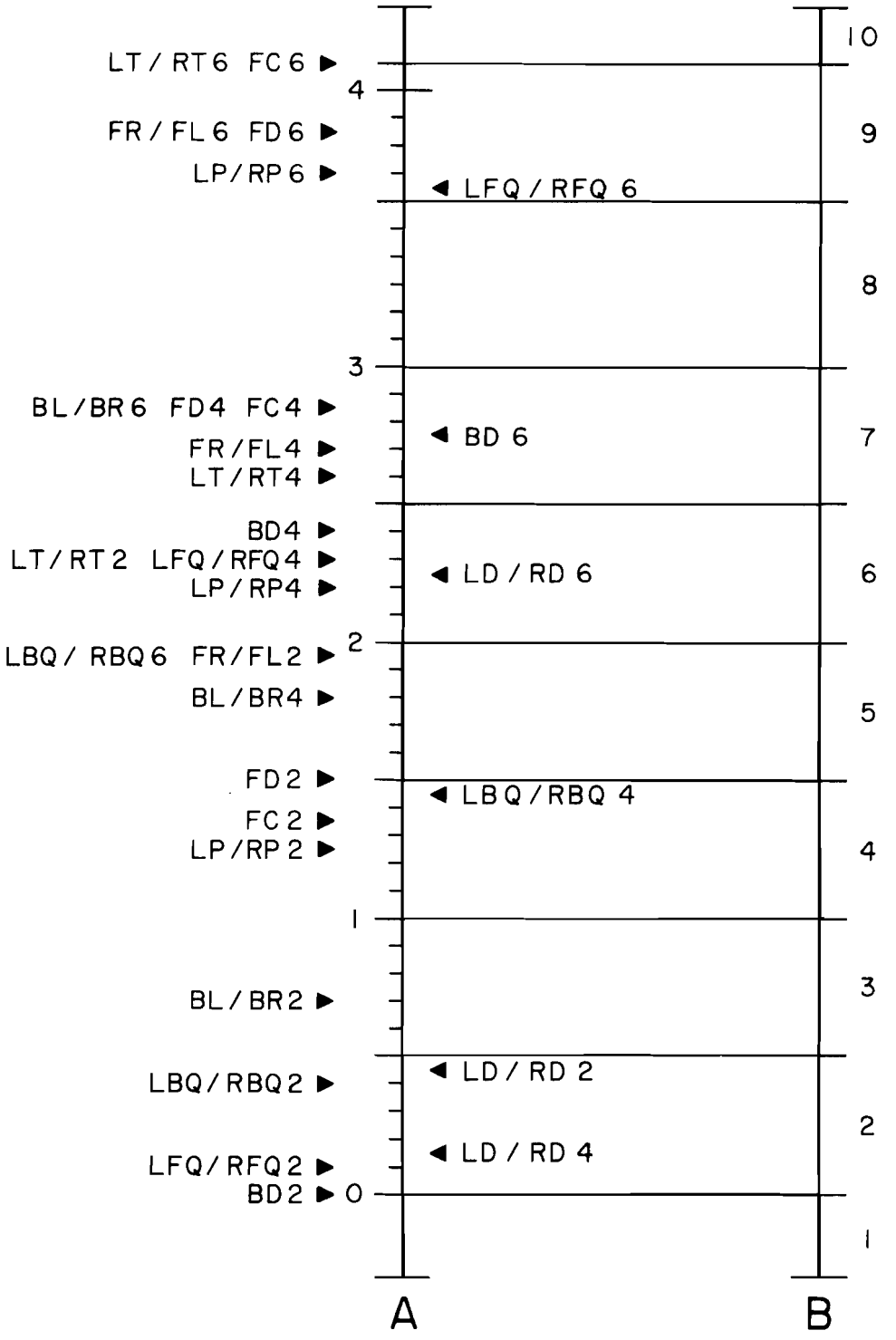

Figure 2. A: Graphic representation of the relative location of each picture of the TAD damage rating scale according to the derived weights shown in Table 3. B: Reassigned severity level on ten point scale for corresponding damage type and severity level shown in scale "A".

TABLE 4

CONVERSION SCHEDULE FOR CHANGING "REPORTED DAMAGE" INTO 10 POINT DAMAGE SCALE

	REPORTED DAMAGE SEVERITY						
DAMAGE TYPE	1	2	3	4	5	6	7
FC	3	4	6	7	8	9	10
FD	3	4	6	7	8	9	10
FL/FR	3	5	6	7	8	9	10
BD	1	2	4	6	7	7	9
BL/BR	2	3	4	5	6	7	9
LP/RP	3	4	5	6	7	9	10
LFQ/RFQ	1	2	4	6	7	9	10
LBQ/RBQ	1	2	3	4	5	5	8
LD/RD	1	2	2	2	4	6	8
L&T/R&T	5	6	6	7	8	9	10

PREDICTION OF DRIVER INJURIES

The objective of developing a good measure of crash severity is to enable the researcher to control for the severity when assessing the crashworthiness of various vehicle designs and to test for the relative value of various vehicular safety devices in terms of injury reduction. Since the TAD damage rating scale requires special training plus the cost of manuals, as well as additional personnel time for coding and keypunching, it is important that it be demonstrated that the scale has a greater potential as a control variable than presently used measures of crash severity. The two measures of crash severity currently available in the North Carolina computerized traffic records system are speed and an estimate of the dollar damage to the vehicle. This part of the report consists of various comparisons between the TAD scale, cost, and speed as predictors of driver injuries, in a sample of 1,329 cases.

PROCEDURE - Copies of the accident reports filed by the Highway Patrolmen who were using the TAD scale were collected by the HSRC and relevant information on each accident vehicle was coded and keypunched.

Speed - During the coding, speed ratings for each vehicle were assigned according to the following system:

> In crashes in which the vehicles were going toward each other and resulted in a front end or sideswipe collision, the speed of the vehicle going the fastest was assigned to both vehicles.
>
> In single vehicle accidents, the speed was assigned directly.
>
> In crashes in which both vehicles were going in the same direction and resulted in a front to rear crash or a sideswipe, the difference in speed between the two vehicles was assigned to both vehicles.
>
> In angle collision, the speed of the striking vehicle was assigned to both the striking and the struck vehicle.
>
> In multiple vehicle collisions, the above rules were followed; however, it was not always possible to assign a speed to all of the vehicles.

Speed, as a variable, was transformed into a seven-point scale for analytic purposes. The transformation was based on the distribution of speeds which were gouped according to the following approximate cumulative percentages: 7, 20, 38, 62, 80, 93, 100, based on the normal curve. This procedure resulted in the speed range groupings on a seven-point scale, as shown in Table 5.

Cost - Experience with cost estimates of accidents has well established the fact that the estimate is to some extent dependent on the age of the vehicle. It was therefore decided to adjust the cost estimates by vehicle age. (It should be noted that vehicle age is not a standard information item on the computerized data file in North Carolina, although it is included in the original accident report. However, plans have been developed to make vehicle age a part of the computerized data file.) Accordingly, a distribution of cost by vehicle age was made, and inspection of this distribution indicated that vehicle age could be reduced to four categories. The four categories of age are 1959 and earlier, 1960-1963, 1964-1966, and 1967 and later. The cost by age category data were then divided into a seven-point scale using the same technique described for categorizing speed data. The results of this grouping are shown in Table 6.

A test of correlation between vehicle age and cost of repair according to the conversion schedule in Table 6 yielded an $r = -.008$ (N.S.) for 1,329 vehicles and indicates that the attempt to correct estimated cost for age of vehicle was successful. (The correlation between vehicle age and unadjusted cost was $r = -.185$.) It should be noted that the range of cost, for each vehicle age group, is widely different due to the fact that the lower boundary for each group approached zero cost, and only

TABLE 5

SCHEDULE FOR CONVERTING SPEED
INTO A SEVEN POINT SCALE

	SCALE VALUE						
	1	2	3	4	5	6	7
SPEED	0-4	5-24	25-39	40-49	50-59	60-69	70+

TABLE 6

CONVERSION SCHEDULE FOR TRANSFORMING COST ESTIMATES
CORRECTED FOR VEHICLE AGE INTO A SEVEN-POINT SCALE

VEH. AGE	SCALE 1	2	3	4	5	6	7
67-69	$ 0-49	50-124	125-249	250-499	500-999	1000-2250	2251+
64-66	$ 0-49	50-124	125-199	200-374	375-699	700-999	1000+
60-63	$ 0-24	25-99	100-149	150-299	300-499	500-699	700+
-59	$ 0-24	25-99	100-174	175-249	250-349	350-499	500+

the upper boundary was more a function of age. This tends to lower a correlation coefficient when unadjusted cost values are used, and, in effect, makes it improper for correlational analysis.

Damage Rating - In many cases the accident reports had more than one damage rating for a vehicle. In these cases only the most severe damage was used for analytic purposes. The most severe damage was selected in two different ways, according to the damage rating scale used (TAD scale or NSc). When the TAD scale was used, the most severe damage was selected by referring to the severity level only, without reference to damage type. When the NSc was used, the most severe damage was selected according to the conversion schedule, shown in Table 4.

Injuries - Since the number of severe injuries in the sample was small, the injury variable was collapsed from a five-point scale into fewer than five points in three different ways for analytic purposes. The five-point injury scale is based on standard injury classifications (4). The derived injury scales and assigned weights are as follows:

Inj. 3 (3 levels)	Inj. 2 (2 levels)	Inj. O/S (2 levels)
0. no injury	0. no injury	0. no injury plus "C" & "B" injuries
1. "C" plus "B" injury	1. any injury plus fatal	1. "A" injuries plus fatals
2. "A" plus fatal		

RESULTS AND DISCUSSION -

Test for Correlation between Injury and Speed, Cost, TAD Scale, and NSc - The results of this test on 1,329 cases are shown in the correlation matrix in Table 7. Column 1 shows the correlation coefficients between variable 1 (Inj. 2) and each of the four predictor variables. All four of these correlation coefficients are significant at the .01 level. However, it will be noted in column 1 that the correlation between speed and Inj. 2

TABLE 7

CORRELATION MATRIX OF INJURY
SCALES AND PREDICTOR VARIABLES

	VARIABLE							
VARIABLE		1	2	3	4	5	6	Inj. O/S
1.	INJ. 2	1.00	.931					
2.	INJ. 3	.931	1.00					
3.	TAD	.433	.451	1.00	.854	.285	.695	.378
4.	NCs	.445	.457	.854	1.00	.304	.659	.375
5.	SPEED	.254	.265	.285	.304	1.00	.338	.223
6.	COST	.443	.431	.695	.659	.338	1.00	.322

% of no injury = .743 % of B or C injury = .133 % of A or K injury = .124
N = 1329

is considerably less than the other predictors. These figures also show that the TAD scale, NSc, and cost are approximately equivalent in predicting Inj. 2. Column 2 shows the correlation between variable 2 (Inj. 3) and the four predictor variables, and it will be noted that the results are about the same as for Inj. 2. However, it should be noted that the cost estimates adjusted for vehicle age represents a unique treatment. The correlation coefficient between unadjusted cost estimates and driver injury for the Inj. 3 category is .303, and this value is significantly less than the value of .431 (Table 7) at the .01 level. Column 7 shows the correlations between Inj. O/S and the four predictors. In this case speed remains a poorer predictor of injury than the other three variables; however, the TAD scale and NSc appear better than cost as predictors, although all are significant at the .01 level.

To test the relative predictive power for the four variables according to injury category, a "t" test for the differences between correlation coefficients (5) was applied. The results of these tests are shown in Table 8. These tests show that for Inj. 2 and Inj. 3 there is no significant difference between the TAD scale, NSc, and cost, but all three are significantly higher than speed (.01 level). The test between the four predictor variables for the Inj. O/S variable shows that no difference exists between the TAD scale and NSc ($t=0.2$), but both are significantly higher than speed ($t=5.1$) and cost ($t=2.8$).

These tests* show that the TAD scale, NSc, and cost are better than speed in predicting driver injuries for each of three injury category groupings. Also, they show that there is no significant difference

---

* These tests are not independent, but by using the .01 level of significance as a reference, one may be confident in a conclusion that the "r" values are significantly different at the .06 level.

TABLE 8

TEST FOR DIFFERENCES BETWEEN r's
FOR PREDICTING DRIVER INJURIES.

INJ. 2

variable

	3	4	5	6
3	/////		**	
4	/////	0.9	6.1	0.5
5	/////	/////	**	
6	/////	/////	6.6	0.1
	/////	/////	/////	**
	/////	/////	/////	6.7

INJ. 3

variable

	3	4	5	6
3	/////		**	
4	/////	0.4	6.4	1.1
5	/////	/////	**	
6	/////	/////	6.7	1.3
	/////	/////	/////	**
	/////	/////	/////	5.9

INJ. O/S

variable

	3	4	5	6
3	/////		**	**
4	/////	0.2	5.1	2.8
5	/////	/////	**	**
6	/////	/////	5.1	2.5
	/////	/////	/////	**
	/////	/////	/////	3.3

** significant at .01 level or more

between the TAD scale and the scale derived from the psychological scaling of the TAD scale into ten points (NSc). The tests also show that cost, due to a special treatment involving adjustments for vehicle age, is as good a predictor of driver injury as the TAD scale when injuries are grouped according to the Inj. 2 and Inj. 3 scheme. The superiority of the TAD scale over cost adjusted for vehicle age appears only when one is concerned with investigating serious injuries (A plus fatal) versus minor injuries plus no injuries (the Inj. 0/S scale).

Multiple Correlation Analysis of Predictor Variables - A multiple correlation test was conducted to determine what gain in predictive power could be achieved when the variables were combined. The results of the multiple correlations are shown in Table 9. The multiple correlation between Inj. 2 and the TAD scale plus speed plus cost was .486. For the same variables the multiple correlation for Inj. 3 was .493. These correlation coefficients are higher than the single variable correlations shown in Table 7, and a test for the differences between correlation coefficients showed that the use of three predictor variables results in a significant increase over the use of the TAD scale or the NSc alone ($F > 30$ with df 2 and 1,325, significant at more than the .01 level).

A second treatment of the data using multiple correlation techniques involved the use of a dummy variable in which two dummy variables were substituted for the Inj. 3 category values 0, 1, and 2 in an attempt to find if different weightings for injury would result in better prediction. The problem of concern here was to find out whether or not numeric weightings for injury severity could best be represented by a simple arithmetic progression (e.g. 0, 1, 2) or by a more complex progression (e.g. 0, 2, 8).

TABLE 9

MULTIPLE CORRELATIONS

		TAD scale	Speed	Cost	R	New scale	Speed	Cost	R
Inj. 2	Beta weights	.233	.104	.246	.486	.257	.094	.242	.495
Inj. 3	Beta weights	.282	.119	.195	.493	.290	.109	.203	.499

The method was to obtain the multiple correlation and canonical correlation between these dummy variables and the TAD scale, speed, and cost or NSc, speed, and cost. The results of this procedure showed a slight but non-significant increase in correlations. This is interpreted as indicating that the three-level injury category with preassigned values of 0, 1, and 2 represents appropriate weightings.

SUMMARY AND CONCLUSIONS

The National Safety Council damage rating scale has been in use in a two-county district in North Carolina for several months. This pilot program is being conducted to determine the usefulness of the scale in describing vehicle damage. Questions about the scale are grouped into three categories and represent three areas of research.

The first research area was concerned with inter-rater reliability. The questions here were: 1) To what extent is the damage rating scale used in the same way, or in different ways, by the Highway Patrolmen presently using the scale? and 2) Are there any tendencies for individual patrolmen to rate vehicles differently from others? This problem was explored by having 17 of the Highway Patrolmen rate each of 25 preselected crash vehicles. The data showed that the men agreed on the type of damage on a vehicle about 80% of the time and that their rating of the severity of damage agreed about 90% of the time. Most discrepancies in damage type could be attributed to slight differences in opinion about the direction of impact on the vehicle, and it was concluded that there was relatively good agreement for the group in regard to the use of the scale in rating vehicle damage. The data also showed that all of the troopers except one tended to evaluate damage type in the same way.

The single exception was a trooper who was on special assignment and only rarely investigated accidents. Further analysis showed that some of the troopers tend to consistently either overestimate or underestimate the severity of damage as compared to the group estimate.

The general conclusion for the study on inter-rater reliability is that there are some problems in using the scale which should not be ignored. However, these problems do not seem to be particularly serious, and there are indications that these problems could be eliminated by the use of a revised manual with better pictures and an improved training program.

The second research area was concerned with the scale characteristics of the various scales in the damage rating manual. The scales were intended to represent equal intervals for each of the six lower points with the seventh point representing damage above a certain level. A psychological scaling study showed that none of the scales had equal appearing intervals and some of the scales had pictures, supposedly showing different levels of damage, which actually appeared to show about the same amount of damage. These findings indicate that the damage rating manual needs improvement and supports one of the conclusions drawn from the study on inter-rater reliability. The psychological scaling technique used in this study would he most appropriate for selecting pictures for a revised manual.

The last area of research concerned the correlation between driver injuries and damage severity ratings in comparison to reported collision speed or estimated dollar damage and driver injuries. The objective was to find the best predictor of injuries for use as a control variable for evaluating vehicle safety design features. Damage ratings, speed, and cost were found to be significantly correlated with driver injuries, but

both damage rating scales and cost were significantly better than speed. Cost was found to be as good as damage ratings in predicting driver injuries in some cases, but when severe injuries are contrasted with minor injuries plus no injury, the damage rating is significantly better than adjusted cost as a predictor. The value of cost estimates in predicting driver injuries was due to a special treatment of the cost in which it was adjusted for the age of the vehicle. This approach could prove valuable in research in which damage ratings are not available.

In general, it seems that the equivalence in predictive power, in some cases, for adjusted cost and damage ratings would disappear as the damage rating scale is improved. This is based on the assumption that damage cost estimates will not improve, or, if they could be improved by special training, an improved damage scale would still be more informative and practical for the researcher. This is asserted on the grounds that the analyses reported here were conducted using only the severity level of damage and discarding the additionally important information regarding damage type. This was done because the sample size was not large enough to allow a breakdown by damage type (17 categories) and severity level (7 categories).

In conclusion, each of the three areas of research reported in this paper indicate that problems exist in regard to the TAD vehicle damage rating scale. However, this should come as a surprise to no one. The TAD damage scale, as it presently exists, was set forth as an experimental tool to be tested and evaluated. Thus, one should consider the present TAD scale as being in only rudimentary form, and it has done admirably well under these circumstances. Accordingly, the research reported here should be viewed as an effort to provide information which would be useful in developing the TAD scale to its fullest potential.

REFERENCES

1. Siegel, Sidney. Nonparametric Statistics - For the Behavioral Sciences. New York: McGraw-Hill, 1956.

2. Bock, R. Darrell and Jones, Lyle V. The Measurement and Prediction of Judgement and Choice. San Francisco: Holden-Day, Inc., 1968.

3. Torgerson, Warren Stanley. Theory and Methods of Scaling. New York: John Wiley & Sons, Inc., 1958.

4. Manual of Uniform Definitions of Motor Vehicle Accidents. USA Standards Institute, New York, New York.

5. Guilford, Joy Paul. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill, 1956.

APPENDIX I

Vehicle Crash Categories and Code Designation

Front End Damage:

FD: Front Distributed; damage extends across front of vehicle.
FC: Front Concentrated; damage due to collision with a narrow object.
FL/FR: Front Left/Front Right; damage restricted to a corner of the vehicle.

Rear End Damage:

BD: Back Distributed; damage extends across rear of vehicle.
BL/BR: Back Left/Back Right; damage restricted to a corner of the vehicle.

Side Damage:

LD/RD: Left Distributed/Right Distributed; damage due to a "sideswipe" collision.
LBQ/RBQ: Left Back Quarter/Right Back Quarter; damage due to an angle impact
LP/RP: Left Passenger Compartment/Right Passenger Compartment; damage due to an angle impact.
LFQ/RFQ: Left Front Quarter/Right Front Quarter; damage due to an angle impact.

Roll Over Damage:

R&T/L&T: Right and Top/Left and Top.