

# THE EFFECTS ON RECALL AND ATTITUDES OF THREE METHODS OF PRESENTING TEST INFORMATION TO 9TH GRADE STUDENTS

John E. Riley and Edgar A. Moorhead  
Florida State University

Counselors in Florida are expected to communicate the results of the state 9th grade testing program to students and parents. This test information is frequently presented using graphs, scores, or a combination of graphs and scores. The purpose of this study was to examine the effectiveness of these three methods of interpreting test information to students.

Although there has been extensive research on standardized test construction, validity, and administration, Stephenson (1964) has noted that in the area of communicating test results there are few published references. A few studies have investigated the problem of test communication. Holmes (1964) examined four methods of explaining test results to college freshmen and found differences attributable to the personality of the counselor rather than the method of explanation. Wright (1963) compared group presentation with individual interpretation of test information and found no differences between the two approaches. Lister and Ohlsen (1962) investigated the effect of orientation on the subsequent communication of test results. They found an increase in the use students made of test information following an orientation to the tests. These studies of the effectiveness of test interpretation seem to have established that test explanation - as one would expect - leads to fairly accurate recall of test results over varying periods of time. There is less information on which methods of presenting test results encourage more accurate recall, and there has been very little investigation of the effects of various methods on the attitudes of students.

## Procedure

Ninth grade students in the laboratory school at Florida State University were randomly assigned to three groups to receive an interpretation of their 9th grade test results. One group (G) received their test results in graphic form. The graph was based on percentiles and indicated an average band between the 40th and 60th percentiles. It provided, however, no scores for student use. A second group (S) received their results in percentile scores. The scores were provided on a slip of paper indicating the content of the tests. The final group (S&G) received both

their percentile scores and a graph. This information was provided on the standard forms for reporting results provided through the state testing program. Each of these groups was further broken into small groupings of six to eight students and randomly assigned to two counselors for group test interpretation.

Prior to receiving their test results, students predicted their test scores and completed a Semantic Differential reporting their attitudes toward various aspects of test interpretation. One week following the test interpretation, students again estimated their test scores and repeated the Semantic Differential. Changes on these instruments were used to assess the effectiveness of the three methods of test interpretation.

Test scores were predicted and then "recalled" on a graph similar to the one used in the graph interpretation. Percentile scores, however, were provided on the instrument. These graphs were scored for absolute differences between estimated scores and the percentiles actually achieved. That is, differences between predicted and achieved scores were obtained on the pre-test and summed for all nine subscores without regard to whether the predictions were higher or lower than actual scores. The same procedure was used to estimate recall following the test interpretation. Effective test interpretation should lead to a decrease in the absolute difference between estimated and achieved test scores.

The Semantic Differential (1964) is a list of bipolar adjectives individually constructed to fit a particular research interest. Adjective pairs - for example, good-bad - are selected from a list on the basis of their factor loadings and their appropriateness to the topics to be investigated. The adjectives are placed on a seven point scale, and students are asked to rate each "concept" on each adjective pair. A "concept" may be a thing, idea, or person. The instrument may yield a number of factors which describe the meaning of a concept. In this study, however, only one factor, Evaluation, was pertinent. Osgood (1964) has suggested that Evaluation is the attitude factor in the Semantic Differential. The following items were used to measure attitudes toward the test interpretation:

<u>Adjective Pair</u>	<u>Evaluation Factor Loading</u>
positive - negative	.48
high - low	.45
beneficial - harmful	.56
kind - cruel	.52
successful - unsuccessful	.51

Students were asked to apply the scales to the following concepts:

9th grade test interpretation	Guidance counselor
My test scores	My ability to do
My achievement in school	school work

In addition to these concepts which occurred in both the pre- and post-testing, students were asked to rate the particular form in which they received their test results on the post-test only. A decline in Semantic Differential scores from pre- to post-test suggests the development of less positive attitudes toward the concept rated.

### Results

The data describing recall of the test results are presented in Table 1. The S and S&G groups significantly improved estimates of their test results following the test interpretation. The G group was not significantly more accurate on the post-test estimate of their test scores. The data did not present information which would be useful in discriminating between S and S&G effectiveness.

Table 1  
Differences Between Median Estimated and  
Median Achieved Test Scores

Group	Pre- test median	Post- test median	Wilcoxon	Signed	Rank-Paired	Replicates
			T	N	Signif.	Direction
Graph	82	66	104.5	22	ns.	-----
Score	111	27	38.0	24	.01	more accurate
Score & graph	57	19	48.0	25	.01	more accurate

As indicated in Table 2 there were changes in attitudes toward aspects of the test interpretation. Members of the G group were less positive in their attitudes toward "9th grade test interpretation" and "Guidance Counselor" after receiving their scores. The S group was less positive toward "Guidance Counselor" and "My own abilities" after the test interpretation.

Semantic Differential scores on the concept "The way I received my test scores" administered after the test interpretation were analyzed using the Wilcoxon Multiple Comparisons Test - One Way Classification (1964). The graph group had a median score of 19.0, the score group a median of 2.0, and the score and graph group a median of 25.5. The G group was significantly lower than the S group at the .05 level and lower than the S&G group at the .01 level of confidence. There was no difference between the S and S&G groups.

#### Discussion

The results of the present study suggest that graphic presentation of test results to 9th grade students was less effective than presenting scores or scores and graphs combined. Following graphic presentation students did not improve their score estimates; attitudes toward the test interpretation became less positive; and attitudes toward the graphic method were negative. Presenting test information graphically appears to have the disadvantages of inaccurate recall and student dislike.

Attitudes toward "Guidance Counselor" become less positive for the G and S groups following the test interpretation, and this seemed to occur for both counselors. In defense of the counselors, however, it may be pointed out that the attitudes of the students were initially very positive, and the decline may be comfortingly interpreted as a reasonable adjustment to realism. The less positive attitude toward "My ability to do school work" in a single group provides too little information to hazard an interpretation.

#### Summary

Florida 9th grade test results were interpreted to students using graphs, scores, or graphs and scores. Pre- and post-testing suggests that graphs were least effective in presenting test information. Students receiving graphed

Table 2

Pre- and Post-Test Semantic Differential Scores:  
Evaluation Factor

Concept	Group	Pre-Test	Post-Test	Wilcoxon T	N	Signed Rank	Paired Replicates
		Mdn	Mdn				
9th Grade Test Interpretation	G	25.5	22.0	33.5	20	.01	less positive
	S	26.0	27.0	72.0	18	---	-----
	S&G	28.0	27.0	99.0	21	---	-----
My Test Scores	G	23.0	20.5	79.5	22	---	-----
	S	24.0	23.0	94.5	22	---	-----
	S&G	24.0	23.5	132.0	23	---	-----
My Achievement in School	G	26.5	26.0	62.0	20	---	-----
	S	27.0	27.5	118.0	22	---	-----
	S&G	26.5	26.0	124.5	24	---	-----
Guidance Counselor	G	30.5	27.5	33.5	20	.01	less positive
	S	30.5	27.5	25.0	21	.01	less positive
	S&G	27.0	26.5	75.0	23	---	-----
My ability to do School work	G	27.0	26.5	56.5	20	---	-----
	S	28.0	27.0	36.0	21	.01	-----
	S&G	27.0	26.5	113.0	24	---	-----

Note--The highest possible score was 35, midpoint 20, and lowest score 5. A positive attitude is indicated by a high score. A

test information did not improve the accuracy of their test score estimates and indicated less positive attitudes toward that method of receiving test information.

#### References

- Holmes, J. The Presentation of Test Information to College Freshman. Journal of Counseling Psychology, 1964, 11, 54-58.
- Lister, J. & Ohlsen, M. The Effects of Orientation to Testing on Motivation for and Outcomes of Test Interpretation. Cooperative Research Project #1344, College of Education, University of Illinois, 1962.
- Osgood, C. E. The Measurement of Meaning. Urbana, Ill., University of Illinois Press, 1964.
- Stephenson, R. Client Interpretation of Tests. Vocational Guidance Quarterly, 1964, 12, 51-56.
- Wilcoxon, F. & Wilcox, R. Some Rapid Approximate Statistical Procedures. Pearl River, N.Y.: Lederle Laboratories, 1964.
- Wright, E. A Comparison of Individual and Multiple Counseling for Test Interpretation. Journal of Counseling Psychology, 1963, 10, 126-134.