

## Procedures and Results of an Evaluation of Writing

Leslie Hendrickson  
Eugene Public Schools, Oregon

### Introduction

Extensive public concern has been expressed in the 1970's about the effectiveness of public school instruction in reading, writing and mathematics (cf. Copperman, 1980; Newsweek, 1975). Despite this concern, no public school district, to my knowledge has carried out a comprehensive, district-wide evaluation of language arts instruction. Two likely reasons exist to account for this. First, the curriculum area known as language arts covers five large subareas: handwriting, language usage, spelling, listening and speaking, and writing. It is difficult to measure all the resources allocated to language arts instruction and the way these resources are used in each subarea in all schools. It is also very difficult to collect large numbers of valid and reliable measures of student ability across all of these areas. Secondly, in the subarea of writing, this difficulty is compounded by the absence of standardized tests and the labor intensive methods used to analyze writing samples.

The Eugene, Oregon school district is the third largest in Oregon: 20,000 students and 1,100 teachers. In the fall of 1977, the Eugene School Board directed its Research, Development and Evaluation (RD&E) unit to look at the district's writing instruction. As one part of its evaluation, RD&E studied writing exercises obtained from a one-third random sample of the district's sixth graders. This paper contributes to evaluation the-

ory by describing the methodology and results of a study of writing and spelling ability. In this paper, I discuss the advantages and disadvantages of two customary methods for analyzing students' writing, show how these methods may be combined and describe an effective way of providing useable evaluation information to principals and teachers. The work described here is only one small part of what a thorough evaluation of language arts instruction would entail.

### Holistic and Primary Trait Analysis

Two customary ways of judging writing exercises are distinguished in the literature: holistic scoring and primary trait analysis. Holistic scoring procedures assign a single score to a written exercise which purports to summarize the overall quality of the writing. Typically, people taking the writing test are asked to write on a given exercise. Two independent raters read each exercise and an average of the two ratings is given to the paper. Exercises which receive very dissimilar ratings are read by a third rater.

The National Assessment of Educational Progress (NAEP) used holistic procedures in their 1969 assessment because in that year these were the only procedures available. State Departments of Education, such as California in 1975 and Oregon in 1978, have found holistic procedures feasible for rating large samples of papers. A summary of state-level work has been published by Wisconsin's De-

partment of Public Instruction (Vicki, 1978). In 1977, the College Entrance Examinations Board (CEEB) reinstated its writing requirement and, for the first time, began holistically scoring the writing exercise.

Mullis (1976) succinctly describes the advantages and disadvantages of holistic procedures. The major advantage is that the method provides a reliable ranking of exercises. Most readers will agree on the sorting of papers into categories. The disadvantages are that the scores can be difficult to interpret. In assigning one overall score to each paper, much information is lost in the summary. Other than a higher overall score, what are the differences between better and average pages? What elementary writing instruction should be changed to improve student writing?

Furthermore, NAEP has found that holistic scoring can be difficult to replicate over time. In 1973, Scott Newcomb of the NAEP staff used basically the same group of readers to rate the same exercise they rated in 1969. While the rank order of the exercises was reasonably invariant, the entire distribution of scores shifted upward. The shift in mean holistic scores was significantly higher. The implications of this finding are that in order to longitudinally compare exercises taken at different times, both sets of papers have to be read by the same raters at the same time.

"Primary trait" scoring procedures are typically used as an alternative to holistic procedures. Primary trait procedures identify specific aspects of writing, ranging from mechanical grammar and punctuation to variables reflecting the exercise's logical organization, form or creativity. The key distinction between holistic and primary trait analysis is that the latter uses multiple variables to characterize the exercise, rather than using one overall global rating of the exercise's quality. The number and variety of variables chosen depends on the purposes and resources of the study, the characteristics of the writer and the ability to develop clear definitions and scoring procedures for the pertinent variables.

Mullis (1976) lays out the rationale for

using primary writing trait analysis and provides examples of its use. Primary trait procedures were used by NAEP in its 1974 writing assessment and by the Washington Department of Education in their 1976 statewide assessment. For a broad overview of writing assessment, and a comparison of holistic and primary trait procedures see Clearinghouse for Applied Performance Testing (1978). These two procedures are often discussed as though they were alternatives to each other. As the following discussion shows, these procedures should not be thought of as mutually exclusive alternatives. Rather, they can be used together in revealing and useful ways.

In recent years, many school districts have used holistic and/or primary trait methods to analyze their students' writing ability. The exact number of districts is difficult to determine. Based on impressionistic comments made by people in other school districts, I estimate that at least 250 districts have done a systematic analysis of students' writing. Some of this work is published, although not in professional journals (e.g., Averill, 1979; Shellenberger, 1978), but most of it is unpublished. For example, the Tigard, Oregon, and Cherry Creek and Westminster, Colorado districts are currently carrying out multiyear efforts to analyze students' writing, but no data from these efforts have yet been published. While some agencies, such as the Northwest Regional Educational Laboratories, have been conducting small studies of school district work, no data bank or annotated bibliography of work currently exists.

### Methodology of Eugene Study

In order to evaluate writing instruction, evaluators first talked with approximately 25 district staff and outside consultants. In April 1978, the Oregon State Department of Education did an holistic analysis of 10,000 exercises and the department's work was observed by Eugene evaluators.

Consultants and district staff recommended that writing samples of sixth-grade students be taken for two reasons. First, in the Eugene district, the sixth

grade is the  
struction i  
results co  
and fed ba  
Moreover,  
to junior h  
junior hig  
these inc  
weak. Sec  
materials  
sixth grad  
It is possi  
tion betwe  
what has l

Two w  
pretested,  
sixth grad  
sample of  
dents.

A comp  
the two e  
ates of Co  
"terms ar  
derstood  
average s  
also show  
length ir  
and 17.0  
grade stu  
the test  
dents. It  
dents he  
cises.

Exerci  
purpose  
familiar  
would  
thoughts  
uators b  
grade t  
1978):

Pre  
who  
the v  
best c  
judge  
to spe  
like t  
would  
that j  
place

Exer  
purpos  
dents l

grade is the culmination of 6 years of instruction in the elementary schools. The results could be summarized by school and fed back to each elementary school. Moreover, pooled results could be given to junior highs, so English teachers in the junior highs would know in what areas these incoming seventh graders were weak. Secondly, the district's curriculum materials contain expectations as to what sixth graders should be capable of doing. It is possible to do a discrepancy evaluation between what should be learned and what has been learned.

Two writing exercises were drafted, pretested, and sent to approximately 500 sixth graders, representing a 30% random sample of each school's sixth-grade students.

A computerized readability analysis of the two exercises by Britton and Associates of Corvallis, Oregon, showed that the "terms are relatively easy and can be understood by the majority of students in an average sixth-grade class." The analysis also showed that the "average sentence length in both exercises is long (22.60 and 17.00 words) for sixth and seventh-grade students." Persons administering the test read each exercise to the students. It is reasonable to infer that all students heard and read the required exercises.

Exercise One is presented below. The purpose of Exercise One was to provide a familiar topic about which all students would be expected to have some thoughts. The topic was selected by evaluators based on suggestions from sixth-grade teachers. Exercise One (May 1978):

Pretend this is a contest. Students who write a good description will get the vacation they describe. Write the best description you can. Make sure the judges know *the place* you would like to spend your vacation, *what* you would like to do when you are there, *why* you would like to do those things, and why that particular place would be a good place to do them.

Exercise Two is presented below. The purpose of this exercise was to see if students knew proper letter formats.

Ned Nelson, who lives at 2360 Geary Street in Eugene, Oregon 97402, ordered 100 special pencils from the Klippo Company for his father's birthday. The pencils were supposed to have *Fred Nelson*, the name of Ned's father, stamped on them.

Ned got the pencils through the mail on April 15, 1978, 3 weeks before his father's birthday. When he unwrapped the package, however, he found that the pencils had been stamped *Fed Nelson* by mistake.

The letter you see on the following page (see below) came with the pencils. Ned decided to write to Mr. Wharton right away and try to get the pencils stamped correctly. On the next page pretend you are Ned and write a *formal* letter to Mr. Wharton.

KLIPPO PENCIL COMPANY  
1040 Clarion Street  
Thompson, Iowa 50301

Dear Customer:

We are happy to fill your order. Did you know that KLIPPO sells more personalized pencils than any other mail-order company? Thanks for making KLIPPO the country's BEST!

We'll be happy to serve you again.

*Sam Wharton*  
PRODUCTS MANAGER

Table I shows the number of students who completed both exercises and the categories of students who did not. Eighty-seven percent of the randomly selected students took the test. There were 435 useable papers written for Exercise One and 428 useable papers written for Exercise Two.

### What Kind of Information About Writing Was Collected?

The first step in the analysis of the exercise was to assign a score to each exercise. Each exercise was read by two independent raters, each of whom would assign a score of 1, 2, 3, or 4, to the exercise. The higher the score, the better the quality of the paper. Five teachers did the rating and the standards they developed and used to assign scores to the exercises are contained in the Appendix to this ar-

TABLE I  
Numbers and Percents of Randomly Selected 1977-78 Sixth-grade Students  
Taking the May 1978 Writing Test

	Number	Percent
Number of students who took the test:	438	87.3
Reasons given to evaluators for not taking the test:		
Moved	17	3.3
Absent	10	2.0
Refused	2	.40
Late submission of test results	5	1.00
Excused for test	4	.80
No reason	26	5.2
Total of Tests Not Taken:	64	12.7
TOTAL:	502	100

ticle. The raters were trained using the same College Entrance Examinations Board procedures used in the State of Oregon's writing assessment. These standards were developed after a review of a random sample of the papers and were further refined during the rating of the papers. These standards established a working definition of what constituted a quality paper. Rating procedures used generally paralleled the procedures used by the Oregon State Department of Education in their 1978 state writing assessment program. Moreover, a rigorous reliability study was made of the raters' internal consistency.

A student's score on each exercise was obtained by adding the two raters' scores and dividing by two. For example, if each rater gave a score of 1 to a paper, the student would get a score of 1.0 for the exercise. If one rater gave a 3 and the other rater gave a 4, then the student's score for that exercise would be 3.5. There were seven possible scores: 1.0, 1.5, 2.0, 2.5, 3.0, 3.5 and 4.0.

Table II shows the raters' scores for the two writing exercises.

The cross-tabulation in Table II shows the results of the two exercises. The data were generated using the Statistical Package for the Social Sciences (SPSS). All of the ordinal correlation coefficients routinely calculated in its cross-tabs subroutines were statistically significant. Students who score well on one exercise

tend to score well on the second. For example, row one shows that the 27 students who received a score of 1.0 on Exercise One also scored either a 1.0, a 1.5, or 2.0 on Exercise Two.

Approximately 40% of the students received a score of 2.0. The average student score on Exercise One was 2.18 and the average student score on Exercise Two was 1.97.

Appendix A contains the standards used by raters to assign scores to students' exercises. Data from Table II can be used in conjunction with the standards to show how many papers met the standards. For example, on Exercise One, 13.6% of the students received a score of 1.5 and 41.8% received a score of 2.0. It is reasonable to infer that the writing of these students was characterized by the standards for a score of 2.0. According to the standards developed by the raters, the writing of over half of the students taking the test was characterized in the following way:

#### Typical Score on 2

The essay only partially addresses the writing task; the purpose or part of the answer is lost. The essay is brief, shows little organization, and demonstrates a lack of detail. There is little evidence of paragraph development. Sentences are typically simple sentences. The essay includes fragments and run-ons. Vocabulary is limited. Spelling, usage, capitalization, and punctuation errors are noticeable.

Sir  
score  
recei  
infer  
writi  
a sco  
dev  
sho  
mo  
thre

Similarly, 16.6% students received a score of 1.5 on Exercise Two and 42.3% received a score of 2.0. It is reasonable to infer that these 58.9% of the students' writing is described by the standards for a score of 2.0. According to the standards developed by the raters, these 58.9% showed "weak performance on one or more of the essential elements." The three elements were correct letter format,

a clear statement of the problem and quality of writing.

These results were reported by the Eugene newspaper and television stations and reporters were quick to publicize them.

### Characteristics of the Writing in Exercise One

In addition to the raters' scores, a pri-

TABLE II  
Raters' Scores for Exercises One and Two—May 1978 Sixth-grade Writing Test

Rater Scores on Exercise One	Rater Scores on Exercise Two								ROW TOTAL
	Count	1.0	1.5	2.0	2.5	3.0	3.5	4.0	
	Row Percent Column Percent Total Cell %								
1.0	21 77.8 32.8 4.8	2 7.4 2.8 0.5	4 14.8 2.2 0.9	0 0.0 0.0 0.0	0 0.0 0.0 0.0	0 0.0 0.0 0.0	0 0.0 0.0 0.0	0 0.0 0.0 0.0	27 6.2
1.5	19 32.2 29.7 4.4	11 18.6 15.3 2.5	23 39.0 12.6 5.3	6 10.2 8.6 1.4	0 0.0 0.0 0.0	0 0.0 0.0 0.0	0 0.0 0.0 0.0	0 0.0 0.0 0.0	59 13.6
2.0	20 11.0 31.3 4.6	34 18.8 47.2 7.9	84 46.4 45.9 19.4	25 13.8 35.7 5.8	13 7.2 44.8 3.0	5 2.8 38.5 1.2	0 0.0 0.0 0.0	0 0.0 0.0 0.0	181 41.8
2.5	2 2.2 3.1 0.5	17 18.7 23.6 3.9	43 47.3 23.5 9.9	20 22.0 28.6 4.6	6 6.6 20.7 1.4	3 3.3 23.1 0.7	0 0.0 0.0 0.0	0 0.0 0.0 0.0	91 21.0
3.0	2 4.0 3.1 0.5	6 12.0 8.3 1.4	22 44.0 12.0 5.1	9 18.0 12.9 2.1	7 14.0 24.1 1.6	4 8.0 30.8 0.9	0 0.0 0.0 0.0	0 0.0 0.0 0.0	50 11.5
3.5	0 0.0 0.0 0.0	2 9.1 2.8 0.5	6 27.3 3.3 1.4	9 40.9 12.9 2.1	3 13.6 10.3 0.7	0 0.0 0.0 0.0	2 9.1 100.0 0.5	0 0.0 0.0 0.5	22 5.1
4.0	0 0.0 0.0 0.0	0 0.0 0.0 0.0	1 33.3 0.5 0.2	1 33.3 1.4 0.2	0 0.0 0.0 0.0	1 33.3 7.7 0.2	0 0.0 0.0 0.0	0 0.0 0.0 0.0	3 0.7
COLUMN TOTAL	64 14.8	72 16.6	183 42.3	70 16.2	29 6.7	13 3.0	2 0.5	433 100.0	

For ex-  
7 stu-  
on Ex-  
a 1.5,  
nts re-  
tudent  
nd the  
e Two  
dards  
o stu-  
II can  
stand-  
et the  
e One,  
core of  
2.0. It  
ting of  
by the  
ling to  
ers, the  
taking  
ollow-  
s the  
f the  
ows  
tes a  
ce of  
s are  
ssay  
ocab-  
ital-  
no-

mary trait analysis was also extracted from the exercises.

Based on recommendations of elementary teachers, seven capitalization, punctuation, usage, and sentence structure variables were chosen for analysis and coded from Exercise One.

Variables were selected, using the following criteria: (1) The skills chosen should be used in the writing exercises collected. For example, according to the district's elementary language arts scope and sequence, the use of the colon in writing time in numerals is to be mastered by the fourth grade. However, very few students use a colon this way in describing vacations they would like to take; (2) The skills chosen should be mastered by the end of and preferably before the sixth grade. For example, if a skill is supposed to have been introduced by the fourth grade, then it is reasonable to assume that a majority of students will be able to perform it by the sixth grade; (3) The skills chosen should be those selected for use in other research studies of writing. For example, skills of capitalization, punctuation, usage, and sentence structure are frequently studied. The seven specific skills given below are in these four general categories; (4) The results chosen should be expressible as rates. For example, the number of sentences in each exercise was coded; thus, rates of error per number of sentences can be calculated.

Additionally, the number of words, sentences and paragraphs in each student's exercise was counted. The results for the first exercise are shown in Table III. The two columns on the left in Table III show the holistic score and number of students whose exercises were coded.

Tables III and IV combine the results of the holistic and primary trait analyses by analyzing the primary trait results across each category of the holistic results.

The next three columns of Table III show that the average number of words, sentences, and paragraphs generally increases as the quality of the writing increases. The seven columns on the right of Table III contain percents.

The number of sentences with proper

capitalization and the number with proper endings both generally increase as the quality of the paper goes up. Overall, nearly 90% of the sentences written had their first word properly capitalized and 89% had a proper ending.

The number of sentences with missing verbs and/or missing subjects neither increase or decrease as the quality of the writing goes up. Ninety-eight percent of the sentences had a subject and almost 100% had a verb.

The last three columns on the right show that there were few sentences with double negatives, two subjects, or misused pronouns.

In summary, capitalization of the first letter of a sentence and sentence ending errors were the most frequent, followed by improper use of pronouns, double subjects and missing subjects. Missing verbs and double negatives are nearly nonexistent.

Three percent of all students taking the test did not complete the final sentence of their paper.

### Characteristics of Writing in Exercise Two

Evaluators also coded seven variables registering the use of parts of a standard letter format: the use and correct placement of the date and zip code, the inside address, the heading, a closing, a greeting, a signature and a body of the letter.

Table IV shows how well students were able to use elements of a letter form. Elements of a letter which were present but incorrectly placed were coded as half missing.

The first two columns on the left of Table IV show the raters' scores and the number of students who received each score. The average score on Exercise Two was 1.97. Students, generally did not do as well on the letter as they did on Exercise One. As with Exercise One, the average number of words, sentences, and paragraphs increases as the quality of the writing increases.

The eight columns on the right of Table IV are expressed as percents of answers that contained the letter element. Approximately 15% of all the letters had a

TABLE III  
Capitalization, Punctuation, Usage, Sentence Structure, and Length of Characteristics  
Exercise One—May 1978 Sixth-grade Writing Test

Rater Score	Number of Students	Words	Average Number of Sentences	Paragraphs	Average Percentage of Correctly Written Sentences									
					Proper sentence ending	Proper capitalization at beginning of sentence	Subjects present	Verbs present	Proper use of single negatives	Proper use of pronouns	Single subjects			
1.0	27	50.5	3.6	1.2	.667	.657	1.000	.963	1.000	.733	.779			
1.5	60	74.0	5.2	1.5	.804	.816	.988	1.000	.994	.893	.943			
2.0	181	86.9	6.1	1.0	.919	.904	1.000	1.000	.980	.963	.973			
2.5	91	114.0	7.8	2.9	.954	.932	1.000	1.000	1.000	.968	.975			
3.0	51	129.0	8.4	3.1	.924	.904	.853	1.000	.997	.994	.994			
3.5	22	134.0	8.6	4.1	.968	.970	1.000	1.000	1.000	1.000	1.000			
4.0	3	262.2	24.3	3.7	.981	.991	1.000	1.000	1.000	1.000	1.000			
Average for all students	2.18	435	97.1	6.7	2.3	.898	.886	.981	.998	.991	.946	.961		

r with  
rease as  
Overall,  
ten had  
ed and

missing  
her in-  
of the  
cent of  
almost

a right  
s with  
r mis-

ie first  
a end-  
it, fol-  
s, dou-  
jects.  
es are

ng the  
itence

ables  
ndard  
place-  
nside  
greet-  
tter.  
lents  
form.  
esent  
s half

f Ta-  
d the  
each  
Two  
ot do  
Exer-  
aver-  
par-  
f the

'able  
wers  
Ap-  
ad a

TABLE IV  
The Letter Form, Exercise Two—May 1978 Sixth-grade Writing Test

Average Percentage of Correctly Written Letter Elements per Exercise											
Number of Students	Average Number of Sentences	Words	Average Number of Paragraphs	Date	Inside address	Heading	Zip code	Closing	Greeting	Signature	Body of letter
64	2.5	44.8	1.2	.620	.620	.203	.352	.484	.789	.797	.969
71	2.9	55.0	1.2	.770	.840	.211	.472	.831	.852	.944	1.000
180	3.3	60.5	1.4	.136	.200	.272	.611	.839	.822	.961	.894
69	3.8	65.8	1.5	.181	.275	.246	.761	.913	.833	.985	1.000
29	3.7	74.5	1.3	.328	.586	.472	.948	.897	.845	1.000	1.000
13	3.2	72.6	1.5	.654	.846	.385	.846	1.000	.961	1.000	1.000
2	5.5	83.0	2.5	.500	1.000	.500	1.000	1.000	1.000	1.000	1.000
Average for all Students	3.3	59.6	1.3	.153	.222	.245	.605	.796	.831	.942	.993

date. A  
had a c  
side ad  
the nan  
ucts m  
cent of  
A corre  
the stre  
code. .  
had a  
closing  
As v  
better  
the err  
Cou  
cause  
questi  
ness f  
Some  
to the  
dents  
was n  
ample  
both  
busin  
Howe  
not c  
resul  
struc




date. Approximately 22% of the letters had a correct inside address. A correct inside address was defined to have either the name of the company, or of the products manager, or both. Twenty-five percent of the letters had a correct heading. A correct heading was defined to include the street address, city and state and zip code. Approximately 60% of all letters had a zip code, approximately 80% had a closing, and 83% had a greeting.

As with the Exercise One results, the better the quality of the paper, the fewer the errors.

Could these results have occurred because students misread or misheard the question, or were confused by the business form letter given in the example? Some confusion might have existed due to the format of the letter. However, students still did not do well where there was no possibility of confusion. For example, the zip codes were included in both the written instructions and in the business letter within Exercise Two. However, 40% of the students' letters did not contain a zip code. Based on these results, it is reasonable to infer that instruction in letter writing has not pro-

vided students an adequate knowledge of how to compose letters of various types.

The elementary language arts scope and sequence of the Eugene School District states that "planned instruction" in writing a "formal, e.g., business" letter shall be given in the fourth and fifth grade and that a knowledge of how to compose letters of "various types using appropriate format" is "essential" for sixth-grade study.

### Spelling

The correct and incorrect versions of all misspelled words in Exercises One and Two were coded. Eighty-four of the 438 students tested made no spelling errors. Out of a total of 67,747 words, 2,663, approximately 4%, or one in every 25 words, was misspelled. It is not known how many of the 67,747 words contained more than three letters. If only words with four or more letters were counted, the rate of misspelling would be higher.

Sixteen words were misspelled 20 or more times. Table V lists these words alphabetically. It gives both the correct and

TABLE V

*The Correct and Incorrect Versions of Misspelled Words and the Number of Times the Words Were Misspelled—Sixth-grade Writing Test May 1978*

Correct Spelling	Number of times Misspelling Occurred	Most Frequent Version of Misspelling	Number of Times Most Frequent Version Occurred
a lot	62	alot	59
because	63	becouse	22
Disneyland	21	Disnyland	10
Hawaii	135	Hawii	105
Hawaiian	21	Hawaian	5
instead	38	insted	25
ordered	23	orderd	14
pencils	64	pensils	10
received	70	recieved	36
sincerely	125	sincerly	71
supposed	20	suposed	4
swimming	22	swiming	15
them	26	then	15
there	56	their	29
Wharton	29	Warton	14
would	22	wound	5

most frequently incorrect version of the words.

The table further shows how often the word was misspelled and how often the most frequent version of misspelling appeared. For example, "a lot" was misspelled 62 times. In 59 of the 62 misspellings, the word was incorrectly spelled as "alot". Exercise One asked students to describe a vacation they would like to take. Disneyland, Hawaii, and Hawaiian were all frequently misspelled.

In Exercise Two, the words "instead", "ordered", "pencils", "received", "suppose", and "Wharton" were also frequently misspelled.

It is reasonable to infer that generally about one to five percent of the students either did not read Exercise Two carefully or did not correctly copy information from it. For example, the word "ordered" appears in line two of the exercise, yet it was misspelled 23 times. The word "pencils" appears seven times in the exercise, yet it was misspelled 64 times. The name of the fictitious products manager, "Mr. Wharton", appears three times in Exercise Two, yet there were 22 misspellings of it.

Another way of showing that some students did not read the exercise is to look at words which were used in the exercise but misspelled. Because almost all students could correctly spell words with four or fewer letters, only words with five or more letters were examined.

Exercise Two contained 45 words with five or more letters. Twenty-eight of these words were misspelled by one or more students even though the correct versions of the words were contained in the exercise.

The word "sincerely" was one of the most frequently misspelled words. It was misspelled by 125 students and there were 32 different misspellings of it. Its frequent and varied misspelling would corroborate the conclusion that most students do not receive instruction in letter writing.

### Distribution of Results

Computer printouts showing the results for each school, compared to regional and district-wide results, were dis-

tributed to each elementary and junior high principal. The distribution was done in small group meetings and included lists of misspelled spelling words. The reaction from the instructional staff was excellent because the data were concrete and easy to understand: for example, 30% of the children from this school did not put a date on their letter.

### Concluding Comments

Holistic and primary trait analyses are not alternatives to each other. Tables III and IV concisely combine the results of the two analyses. The two approaches are only alternatives in the sense that with limited resources of time and money, not all types of analyses can be undertaken.

The number of words, sentences and paragraphs in the exercise had a high linear relationship to the raters' scores. The relationship was so strong that it is reasonable to speculate if rater scores are necessary at all. Multiple regression techniques could be used to assign exercises to meaningful categories using the word, sentence and paragraph length of the exercise.

The advantages of this approach would be a substantial savings because raters would not be required. The disadvantages are that this procedure makes the results harder to interpret to other people. An extensive public discussion followed the release of the writing evaluation results. The results were publicized in the local newspaper and on television and further discussed on two radio talk shows. The holistic results which compared the standards used by the raters to the percentage of students who achieved certain scores were understandable to noneducators.

Data from the primary trait analysis were useful to elementary and junior high principals and junior high language arts teachers because specific areas that needed work, such as letter writing, were identified for each school's students.

The technology of evaluating writing exercises is still primitive. Little theory exists and comparative data are difficult to obtain. There currently exist no data banks showing writing analysis results. For example, although the College En-

trans  
quire  
1960  
resul  
to th  
eval  
1980  
ucat  
distr  
tions

Refe

1. Av  
re  
di  
2. SH  
p  
S

Refe

COPPE  
Q

trance Examinations Board (CEEB) required writing exercises or exams from 1960 to 1970, it did no scoring of the results. The results were simply returned to the schools. A better technology for evaluating writing should emerge in the 1980's as the National Assessment for Educational Progress (NAEP) and school districts publish more of their evaluations.

### Reference Notes

1. AVERILL, M. *Boulder Valley writing assessment report*. Unpublished report. Boulder, Co.: Boulder Valley School District, March 1979.
2. SHELLENBERGER, J. *Composition assessment*. Unpublished report. Beaverton, Ore.: Beaverton School District, October 1978.

### References

COPPERMAN, P. *The literary hoax*. New York: Morrow Quill Paperback, 1980.

- FREDERICK, V. *Writing assessment research report: A national survey*. Madison, Wis.: Wisconsin Department of Public Instruction, May 1979.
- Why Johnny can't write. *Newsweek*, December 8, 1975, 58.
- MULLIS, I. U.S. *The primary trait system for scoring writing tasks*. Denver, Co.: National Assessment of Educational Progress, Education Commission of the States, 1976.
- Clearinghouse for Applied Performance Testing. *Newsletter* (Vol. 3, No. 3). Portland, Ore.: Northwest Regional Educational Laboratory, January 1978.

---

### Author

LESLIE HENDRICKSON, Evaluation Specialist, Research Development and Evaluation Division, Eugene Public Schools, 200 N. Monroe St., Eugene, OR 97402. *Specializations: Methodology*.

(See Appendix on next page)