DOCUMENT RESUME

ED 029 696 PS 001 792

By-Biesbrock, Edieann

A Study of Composition Ability as Assessed With a Standardized Instrument for Second and Third Grade Children.

Pub Date 69
Note-12p.
EDRS Price MF-$0.25 HC-$0.70
Descriptors-Age Differences, *Composition Skills (Literary), *Elementary School Students, Essays, Grade 2, Grade 3, Intelligence Quotient, Measurement Instruments, Rating Scales, Reading Achievement, Sex Differences, Test Reliability, *Test Validity, Vocabulary Development
Identifiers-Global Essay Instrument

Written compositions of 200 second and third grade children were compared four times over a 2-year period to determine relationships of composition ability with intelligence and reading level. A global essay instrument developed at the University of Georgia was used for the comparisons to find out whether or not the instrument revealed any improvement in the essays. Four raters scored each essay independently and without knowledge of the other three ratings by following standardized procedures. Four mean test scores for each student formed a chronological sequence from the fall of 1966 through the spring of 1968. There was a steady increase in the global essay means of both boys and girls over 2 years. Correlations between reading level, vocabulary, intelligence, and global essay scores showed positive relationships. There was not a significant correlation between global essay scores and sex or grade level. The study confirmed the sensitivity of the global essay instrument in measuring the growth trend of second and third grade children. (DO)

A STUDY OF COMPOSITION ABILITY AS ASSESSED WITH A STANDARDIZED INSTRUMENT FOR SECOND AND THIRD GRADE CHILDREN

Edieann Biesbrock, University of Minnesota

Introduction

A global essay instrument[1] developed at the University of Georgia was used to compare written compositions produced at four testing dates over a two-year period. The chief purpose was to discover if the instrument could be used to reveal improvement in the quality of writing produced by the children in the sample group. A second purpose was to examine and identify any differences in growth in composition ability between the boys and the girls in the sample. Finally, a further purpose was to examine the relationships between composition ability and intelligence and between composition ability and reading level, as ascertained with standardized test instruments.

From the results of other research, notably Diederich (1966) and Wilhelms (1967), one can assume that for large groups of children writing does improve from year to year. For the purposes of the study, two-hundred children subjects (one-hundred boys and one-hundred girls) in second and third grades were identified and tested with the global essay instrument. To control for the problems of measuring change, such as fluctuation from testing to testing that is not true growth, several writing samples were taken over specific

---

[1] Biesbrock, E. F. "The Development and Use of a Standardized Instrument for Measuring Composition Ability in Young Children (Grade two and three)." Unpublished dissertation, University of Georgia, 1968. This study was supported, in part, by the English Curriculum Study Center (USOE) and the Research & Development Center (USOE) of the University of Georgia.

intervals during a two year period. Essays were identified and rated from fall 1966, spring 1967, fall 1967, and spring 1968. As recommended by Diederich (1966), these four essays were used to determine a trend of study change in the global essay ratings.

The Test Instrument

The product-scale essay instrument with five forms (essay topics) was developed from a population of more than 1,000 second and third graders.[2] This typical product-scale instrument evaluates the "product" (i.e., the child's written composition) by comparing it to other "products" to derive a relative measure of its merit. Actual second and third grade writing, produced and selected under standardized conditions, served as the models for a seven-point scale. Rater reliability coefficients for the global essay instrument ranged from the mid .60's to the upper .70's, as did the reliability coefficients for equivalent forms. Examinee reliability was checked at .70. Validity checks involving the original criteria for the essay instrument and teacher estimates of student composition ability yielded moderately high coefficients for this kind of test (for validity checks by original criteria .70; for validity checks by teacher estimates .56).

The Sample Population

Two-hundred boys and girls were identified from the schools cooperating with the English Curriculum Study Center at the University of Georgia. This group was composed of fifty second grade boys,

---

[2]Test development reported at 1967 and 1968 NCME meeting by L. R. Veal.

fifty second grade girls, fifty third grade boys, and fifty third grade girls. The subjects were followed over a two year period of development and tested on four separate occasions (each fall and spring for the 1966-67 and 1967-68 school years). The subjects were tested further to obtain subject data of several kinds. The subject variables examined included sex, grade, chronological age, and test scores on reading and mental maturity instruments. ~~Two standardized instruments were used~~.

Third grades received the "California Short-Form Test of Mental Maturity" (1968 revision with 1965 norms; S-Form, Level 1H), and the Reading Test of the "California Achievement Tests" (1957 editions, WXYZ series, 1963 norms; Form W, Upper Primary). The second grades received the "California Short-Form Test of Mental Maturity" (1963 revision, 1964 norms; S-Form, Level 1), and the Reading Test of the "California Achievement Tests" (1957 edition, WXYZ series, 1963 norms; Form W, Lower Primary). The means and standard deviations for chronological age, reading level and vocabulary scores (G.P.'s) from the "California Reading Test" and mental maturity score (I.Q.'s) from the "California Short-Form Test of Mental Maturity" are reported in Table I.

The Rating Procedures

As each essay was collected and processed, it was scored by four trained raters following standardized procedures specified by the global essay instrument. Each rater scored the essays independently and without knowledge of the other three ratings. The four raters revealed moderately high inter-rater reliability over the period of the study (.70). From the four separate ratings, a mean

rating was derived for each essay produced on each of the four testing dates; consequently, for each of the two-hundred subjects there were four mean test scores forming a chronological sequence from fall 1966 through spring 1968. Means, standard deviations, and standard errors for boys and girls for each of the four test dates are reported in Table II. The data were analyzed statistically by means of a two by four factorial ANOV (Table III) and Duncan's Multiple Range Test (Table IV).

Findings

The two by four (sex-by-testing dates) factorial arrangement revealed the testing dates to be highly significant. Sex and the sex-by-testing dates interaction were non-significant. The means of both the boys and the girls for the first testing were significantly different from all subsequent testings, although there were no significant differences between the boys and girls on any single testing or on overall performance. A non-significant change in rate was revealed. There was, however, a steady increase in the means of both boys and girls over the four testings during the two-year period. The combined means for both boys and girls also revealed a steady increase: 3.28, 4.58, 4.69, and 4.82.

It is interesting to note that the largest increase (from the first to the second testing is from 3.28, a quality rating below the mean of the global essay scale (and, consequently, below average quality), to 4.58, a quality rating above the mean. After rising above the mean, both boys and girls advanced more slowly on the quality scale. The tendency to level off above the mean is evident in Figure 1. At all testings, the means of the girls were slightly

higher than the means of the boys, when examined collectively by sex or when sub-grouped by grade.

Additional analysis of subject variables was provided by computing product moment correlations between reading level scores, vocabulary scores of the reading test, intelligence (I.Q.) scores and global essay scores. These correlations showed positive relationships and are reported in Table V. Reading level correlated with essay scores significantly (.65), as did intelligence (.48). The highest correlation between a subject variable and essay scores was revealed between global essay ratings and the vocabulary section of the reading test (.68). Vocabulary has been considered a major component of composition skill in the elementary grades by other researchers. The data in this study support the findings of McLean (1964), Carlson (1963), Howell (1956), and others, namely, in that the children who scored high on the vocabulary measure employed in this study also tended to score high on the global essay instrument.

Chronological age showed little relationship, and that negative, to global essay scores (Table V). While age is considered to be associated with composition quality generally, it should be noted that these four groups had extremely limited ranges. The overall mean of the sample for both grades was 97.74 months with a standard deviation of 7.66. These children were in the second or third grade classes on the basis of common age and/or maturity as discerned within their respective schools. This may explain, in part, the lack of a significant correlation for any group when examined by sex and grade level. However, the correlation of chronological age and global essay quality for the entire sample became slightly

higher (.22), when the range was expanded to include both second and third grade boys and girls.

Summary

The growth study revealed a steady increase in global essay means for the written compositions produced by the children on four testing dates (fall 1966, spring 1967, fall 1967, and spring 1968). This finding is in agreement with Diederich (1966) and Wilhelms (1967) who have reported that growth can be measured for large groups from year to year. It is also a confirmation of the sensitivity of the global essay instrument in measuring the growth trend of second and third grade children. Although the mean essay ratings for girls were slightly higher than for boys consistently for all the testings, both boys and girls showed significant growth during the two year period. Consequently, it can be concluded that, for the children of this study, the rate of growth was not related to sex.

Edieann Biesbrock
University of Minnesota

NCME
Los Angeles, 1969

References

Carlson, R. K. "Recent Research in Originality," Elementary English XL (October, 1963), 583-89.

Diederich, P. B. "How to Measure Growth in Writing Ability," English Journal, LV (April, 1966), 435-499.

Howell, M. "Differentiating Variables in Compositions of Seven Year Olds," Elementary School Journal, LVII (December, 1956), 145-49.

McLean, H. W. "A Comparison of Selected Aspects of Oral and Written Language of Fourth, Fifth, and Sixth Grade Students." Unpublished dissertation, State University of Iowa, 1964.

Wilhelms, F. T. (ed.) "Cooperative Evaluation of English," Evaluation as Feedback and Guide, 1967 Yearbook. Washington: ASCD, 1967.

TABLE I

MEANS AND STANDARD DEVIATIONS FOR
CHRONOLOGICAL AGE, VOCABULARY, READING LEVEL, AND
INTELLIGENCE SCORES

Sex	Grade	No.	C.A. in Months	S.D.	Vocabulary G.P.	Vocabulary S.D.	Reading G.P.	Reading S.D.	Mental Maturity I.Q.	Mental Maturity S.D.
Boy	2	50	92.63	5.80	2.77	.65	2.79	.69	110.61	10.34
Girl	2	50	93.28	4.97	3.02	.58	3.09	.61	109.30	10.71
Boy	3	50	103.29	5.03	4.09	.71	3.96	.56	113.71	10.58
Girl	3	50	104.47	5.77	4.24	.57	4.13	.47	114.86	12.12

TABLE II

MEANS, STANDARD DEVIATIONS, AND STANDARD ERRORS FOR BOYS AND GIRLS ON GLOBAL ESSAY RATINGS FROM FOUR TEST DATES

Sex	No.	Test Interval #1 $\bar{X}$	S.D.	S.E.	Test Interval #2 $\bar{X}$	S.D.	S.E.	Test Interval #3 $\bar{X}$	S.D.	S.E.	Test Interval #4 $\bar{X}$	S.D.	S.E.
Boys	100	3.14	1.02	.15	4.53	1.26	.19	4.55	1.10	.17	4.83	1.11	.17
Girls	100	3.38	.86	.12	4.62	.77	.10	4.80	1.01	.14	4.81	.76	.10
Total	200	3.28	.94	.09	4.58	1.01	.10	4.69	1.05	.11	4.82	.92	.09

TABLE III

ANOV FOR SEX AND TEST DATES

Source	df	SS	MS	F
Total	399	574.21		
Treatments	7	168.66		
Sex	1	.55	.55	.71
Testings	3	167.80	55.93	71.71**
S x T	3	.31	.10	.13
Error	392	305.55	.78	

** Significant at the .01 level.

TABLE IV

COMPARISON OF ESSAY MEANS BY GRADE AND SEX

Comparisons:	B1	G1	B2	G2	B3	G3	B4	G4
Means:	3.16	3.29	4.53	4.58	4.58	4.71	4.87	4.88

[a]

[a]Means not underscored by the same line are significantly different as determined by Duncan's Multiple Range Test at the .01 level.

TABLE V

SUBJECT VARIABLES

CORRELATED WITH GLOBAL ESSAY RATINGS

Sex	Grade	No.	C.A.	Vocabulary G.P.	Reading G.P.	Mental Maturity I.Q.
Boy	2	50	-.26	.55	.55	.31
Girl	2	50	-.18	.61	.68	.48
Boy	3	50	.01	.65	.67	.62
Girl	3	50	-.06	.61	.50	.51
Total		200	-.12	.68	.65	.48

FIGURE 1

GROWTH MEANS FOR BOYS AND GIRLS
FOR FOUR TESTING DATES