DOCUMENT RESUME

ED 038 184 PS 003 164

AUTHOR Sause, Edwin F.; Crowley, Francis J.
TITLE The Development of a Computer Technique for the Content Analysis of Psycho-social Factors in the Oral Language of Kindergarten Children.
PUB DATE Mar 70
NOTE 10p.; Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, Minnesota, March, 1970

EDRS PRICE EDRS Price MF-$0.25 HC-$0.60
DESCRIPTORS Content Analysis, Kindergarten Children, *Language Patterns, Language Usage, *Measurement Instruments, *Measurement Techniques, *Psychological Evaluation, *Sex Differences, Test Reliability, Vocabulary, Word Frequency

ABSTRACT

To find out if a computer could be programmed to efficiently analyze the psychosocial factors in the speech of children, taped language samples were collected from structured interviews with 81 male and 63 female kindergarten children. Thirty psychological and social factors relevant to children's speech were drawn from the words in the samples. A group of school psychologists then independently placed each sample word into its appropriate category to form a content analysis dictionary. Subsequent computer content analysis of data agreed closely with the content analysis performed by a kindergarten teacher. Computer content analysis of sex differences also agreed closely with research and opinion about personality differences in boys and girls. It was concluded that computer analysis of psychosocial factors in the language of young children is a quick, efficient way to gather information that was previously expensive and time consuming to obtain. The second part of this document is a paper which describes the development of the analysis technique used in the language study and contains the data and implications of the project. (MH)

The Development of a Computer Technique for the Content Analysis of Psycho-social Factors in the Oral Language of Kindergarten Children

Edwin F. Sause
Research Associate
Educational Research Council of America
Cleveland, Ohio

Francis J. Crowley
Fordham University
New York, New York

American Educational Research Association
Minneapolis, Minnesota
1970

The measurement of academic achievement of young children is a particular problem for educators. For children who can read, achievement can be evaluated by group tests. But, individualized achievement tests are usually required for non-readers. These tests require the administration and scoring skills of highly trained examiners - whose fees are often expensive. In addition to the expense, achievement tests for pre-readers frequently yield highly subjective measurements.

The trend to place children in school at an early age has aggravated the problem of evaluating the child's progress in the early years. Since the child cannot yet read or write, all measurements, either academic or personal and psychological, are of necessity individualized.

The major channel through which a pre-school child communicates is oral language. If his achievement in attaining both cognitive and non-cognitive goals of curricula is to be assessed, it must be through this means of communication. Ideally, this could be done with a minimum of expense and a maximum of objectivity, reliability, and validity. It is hoped that the present study offers a technique by which this means of assessment could be implemented.

The investigator conducted an inquiry into the nature of the psychological and social factors in oral language as an initial step in the development of an instrument to measure the change of psychological and social factors in groups of children. The technique could also be used to evaluate the effectiveness of pre-school programs designed to bring about changes in behavior indicated by these factors. The assumption that verbal behavior indicates internal psychological conditions is based on previous research in the content

analysis of oral language (Weintraub & Aronson, 1962):

Subjects

The sample for this study was a transcription of a tape recording of the oral language of 81 boys and 63 girls randomly selected from the kindergarten population of Ithaca, New York, in the fall of 1968.

Procedures

A male interviewer recorded an oral language sample for each subject in a standardized situation. An abstract block and a toy fire engine were used as stimuli to elicit verbal responses from the children. The child was asked to tell the examiner everything he could about the stimulus. If the child failed to respond, a series of standardized questions was asked in a specific order.

The 144 interviews were transcribed, and all the different words used by the children were listed out of context. In the first phase of the analysis, the investigator and each of four psychologists examined the list of words in relation to content analysis categories from the Harvard III Psychosociological Dictionary (Stone et al., 1966) and independently chose categories they considered appropriate for analysis of children's language. The 26 categories listed in Table 1 (numbers 1 to 26) were selected because they promised to provide useful information about the expression of psychological and social variables in the language of children.

(Insert Table 1 About Here)

Ten New York City school psychologists, who volunteered for the study because of their interest in children's language, participated in the second

Table 1

The Relative Frequencies and Psychological and Social Tags

	Tag	f-Boys	Index No.	f-Girls	No. of Words Possessing a Tag
1	neutral word	3872	1.6	2371	249
2	aggression	58	6.5	9	20
3	a passive state	1	-	0	1
4	self, (egocentric)	325	1.9	165	19
5	other people	457	1.7	263	16
6	the future	0	-	4	1
7	physical and bodily needs	78	1.9	41	18
8	time	240	2.0	121	26
9	space	446	2.1	209	29
10	quantity	276	2.3	115	30
11	ideational process	286	1.6	172	20
12	affection	0	-	0	0
13	fears	4	2.0	2	6
14	good	15	5.0	3	5
15	bad	16	16.0	1	5
16	superego words or authority, "ought words"	9	2.3	4	6
17	act of oral communication	56	2.9	19	19
18	effort and activity	128	1.7	75	20
19	physical movement	349	2.1	162	68
20	recreation	111	1.6	69	25
21	family and home environment	132	1.6	80	36
22	male	91	1.2	70	10
23	female	9	.5	18	8
24	negative words	169	2.0	84	19
25	affirmation	47	2.9	16	6
26	animals	17	2.1	8	21
27	questions asked the examiner	84	2.9	29	*
28	questions asked by the examiner	642	1.1	599	P
29	indistinguishable sounds	191	2.2	85	3
30	sounds	233	1.6	147	7

phase of the analysis. In order to reduce the fatigue factor, the pages of the list of words were arranged according to a generated table of random numbers and split into halves. Each psychologist received one-half of a list of words, which he was instructed to assign to as many as three of the 26 categories. Each word received five judgments as to its psycho-social meaning. If three of the five judges agreed the word was representative of a category, the word was so tagged and included in the dictionary. Before the final computer analysis, the investigator included four more variables he judged important in the language of children (Table 1, numbers 27-30).

The resulting dictionary was used to study the sex differences in the language of children. The computer program used in this comparison calculated the mean length of response and the type-token ratio for each group as well as the relative frequencies of the 30 content analysis categories.

The concurrent validity of the procedure of assigning tags to words out of context for future analysis of material in context was established by another judge who had not previously participated in the study. The judge, a kindergarten teacher with seven years of teaching experience, analyzed the content of a language sample of ten randomly selected subjects. She was given a list of the 30 categories to be used in analysis along with instructions to count the number of times each category appeared in the text. The procedure was repeated with the same sample after a six hour rest period. A Pearson product-moment coefficient of correlation was computed to determine the scorer's

test-retest reliability.

The same sample of language was analyzed by the computer using the dictionary compiled in this study. The intercorrelations for the three judgments on the 30 categories (N=30) were computed to determine the amount of agreement between the human judge and the machine.

Results

Of the 747 words submitted to the judges, only 22 were found to represent two psycho-social variables. The third column in Table 1 lists the frequency of occurrence of each tag in the dictionary The symbols next to categories 27 and 28 resulted from the peculiar nature of these categories. These symbols were used in the text to indicate the occurrence of a question. The frequency of occurrence was the sum of the frequency for boys and girls.

The relative frequencies for boys and girls in all the categories are listed in columns 1 and 3 of Table 1.

A Spearman rank order coefficent of correlation was computed for the two groups using the categories as subjects. It was found to be .77. This indicated that differences did exist between boys and girls with regard to frequency of references to categories. In order to determine the categories which caused the reduction of the Spearman rank order coefficient of correlation, a ratio number was computed for the number of boys' references to a category as compared to girls' references to the same category. If it differed from the ratio of 1.6 for the entire sample by an absolute value of .5 it was considered to have reduced the Spearman r. This ratio, reported as the index number in Table 1, indicated that boys had proportionately more references to aggression, space, quantity,

good, bad, authority, physical movement, and affirmation and asked the examiner more questions than the girls did. The girls had more references to the female role and required more questions by the examiner to elicit the language sample. Finally, boys made more indistinguishable sounds than the girls made. The psychological significance of these differences was discussed by Sause and Gunning (1969) and Sause (1970).

Table 2 shows the results of a t test for the significance of the difference between the mean word length of independent samples and a t test for the significance of the difference between proportions.

(Insert Table 2 About Here)

The results indicated no significant sex differences in the length or variety of response for boys and girls. This result did not concur with the general opinion of researchers in the field (McCarthy, 1953).

Table 3 lists the coefficients of correlation for the test-retest reliability of the human scorer and the concurrent validity of the computer content analysis technique developed in this study with the hand content analysis.

(Insert Table 3 About Here)

Although the heterogeneity of the scores brings the coefficients of correlation very near the perfect correlation of 1.000, it was found that the human judge working with words in context and the computer working with empirically tagged words out of context scored the frequency of occurrence of the 30 content analysis categories identically.

Table 2

The Results of Tests of Significant Differences between Boys and Girls in Mean Length of Response and Type-Token Ratio

	N	$\bar{X}$ No. of Words	S. D.	t for Words	TTR	t for TTR
Boys	81	204.1	133.1	1.87	.13	.33
Girls	63	163.2	125.8		.10	

Table 3

Coefficients of Correlation for the Relationships Between Computer and Hand Content Analysis

			(1)	(2)	(3)
Human Scorer	first trial	(1)	1.000	.990	.976
Human Scorer	second trial	(2)		1.000	.974
Computer		(3)			1.000

Conclusions and Discussions

It was concluded that a computer can use an empirically constructed dictionary to perform rapid content analysis of words out of context almost as well as a human judge can working with the categories and the words in context

The difference observed in the psychological and social factors present in the oral language of boys and girls seemed to concur with the roles expected of the two groups in western society.

The language indicated that the boys were more aggressive and more interested in evaluation of behavior (good, bad, authority). This followed from the fact that more aggressive people are often subjected to more active restrictions on their behavior by society than the passive conformist. In order to avoid conflict the boys wanted to know more about the situation so they asked more questions and made more affirmations than the girls did. As was expected, the boys were also more interested in quantity, space, physical activity, and animals.

The girls, on the other hand, were more passive, made more references to the female role, and required more questions to elicite verbal response than the boys did.

The conclusion of this study was that it was possible to identify by compu content analysis factors in the verbal behavior of groups of kindergarten childre The information gained by this procedure was previously only available by very slow, very expensive projective techniques. It was now possible to measure behavioral change in groups of students by the content analysis of samples of their verbal behavior. These procedures used equipment available to many researchers in the area of child development.

RECOMMENDATIONS

It is recommended that:

1. These techniques be used in a cross-sectional or longitudinal study of the development of psychological and social factors in the language of children.

2. Other investigators examining psychological and social variables use this technique to test their hypotheses.

3. These techniques be used to evaluate the effectiveness of non-cognitive objectives of curriculum materials.

References

McCarthy, D. A. Some possible explanations of sex differences in language level and disorder. Journal of Psychology, 1953, 35, 155-160.

Sause, E. F. & Gunning, A. A study of sex differences in the oral language of kindergarten children. Paper presented at the Educational Research Association of New York State Convention, Kiamesha Lake, N. Y. November, 1969.

Sause, E. F. Computer content analysis of psychological and social factors in the oral language of kindergarten children. Unpublished doctoral dissertation. Fordham University, 1970.

Stone, P. J. et al. The General Inquirer: A computer approach to content analysis. Cambridge, Massachusetts: M. I. T. Press, 1966.

Weintraub, W. & Aronson, H. The application of verbal behavior analysis to the study of psychological defense mechanisms: Methodology and preliminary report. Journal of Nervous Mental Disorders, 1962, 134, 169-181.