DOCUMENT RESUME

ED 049 590 40 EC 032 209

AUTHOR	Walker, Hill M.
TITLE	Assessment and Treatment of Deviant Behavior in Children - Section Two: Construction and Validation of a Three Stage Model for Identifying and Assessing Deviant Behavior in Children. Final Report.
INSTITUTION	Oregon Univ., Eugene.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
BUREAU NO	BR-6-1308
PUB DATE	67
GRANT	OEG-4-6-061308-0571
NOTE	72p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Behavior Change, *Behavior Rating Scales, Check Lists, Diagnostic Tests, *Emotionally Disturbed, *Exceptional Child Research, *Identification, Models, Research Methodology, *Socially Deviant Behavior, Test Construction, Test Reliability, Test Validity

ABSTRACT

The document, part two of a six part project report, discusses the construction and validation of a three stage model for assessing deviant behavior in children. The model was developed to meet the project's measurement, identification and diagnostic goals. Stage one consists of a 50 item behavior checklist which was used as an initial screening device. The scale is normed on elementary grade subjects; and item reliability and validity estimations are completed on the scale. Stage two consists of a 62 item behavior rating scale which is divided evenly between items to measure acting out behavior and items measuring withdrawn behavior. A behavioral observation form was developed for the purpose of measuring task oriented behavior. The form allows simultaneous observation of 13 behaviors, and contains codes for classroom setting, the social consequences of child behavior, and the social agent supplying the consequence. Identification data were collected in stage one and more specific data were collected in stages two and three for diagnostic and treatment prescription purposes. Further details on the project are contained in section one (EC 032 208) overview; section three (EC 032 210) treatment; section four (EC 032 211) generalization and maintenance; section five (EC 032 212) teacher behavior; and section six (EC 032 213) single subject experiments. (CD)

Final Report

Section Two: Construction and Validation of a Three Stage Model for Identifying and Assessing Deviant Behavior in Children

Assessment and Treatment of
Deviant Behavior in Children

U. S. O. E. Contract OEG 4-6-061308-0571

Hill M. Walker - Project Co-Director

Stage One

Behavior Checklist (WPBIC)

Manual for the

Walker Problem Behavior Identification Checklist

by

Hill M. Walker, Ph. D.

Hill M. Walker
Dept. of Special Education
College of Education
University of Oregon 97403
Eugene, Oregon

Introduction

This manual describes the construction, validation, and administration-scoring procedures of a behavior checklist for the identification of children with behavior problems. The WPBIC is designed for use by the elementary teacher in grades one through seven and is composed of observable, operational statements about classroom behavior which were furnished by a representative sample of elementary school teachers. The checklist is to be used as a supplement in the total identification process rather than as an instrument to simply classify children as emotionally disturbed or socially maladjusted. The WPBIC should function as a tool which the elementary teacher can rely upon in the difficult task of selecting children with behavior problems who should be referred for further psychological evaluation, referral, and treatment.

The Raters

The classroom teacher is in a unique position to identify children with behavior problems since she spends more time in actual observation of the child than any other school personnel. Research studies have demonstrated that teachers are capable of making valid judgements about classroom behavior (e.g. Stouffer,1952; Bower,1958; Beilin,1959; Kaes, 1966). The WPBIC consists of stimulus items which describe behaviors that interfere or actively compete with successful academic performance. The scale is thus especially suited for classroom teachers since according to Beilin (1959), teachers are most concerned with classroom behavior which is disruptive of achievement. Since the teacher is held responsible for the child's achievement through the teaching-learning process, she should be an excellent judge of classroom behavior which is incompatible with academic performance. The teacher is thus regarded as the most qualified rater in using the WPBIC to identify children with behavior problems who are in need of special educational-psychological services. However, ratings from other educational specialists such as counselors, remedial teachers, and school psychologists, who have worked directly with the child, can be obtained for purposes of comparative analysis.

The Period of Observation

A two month observation period should precede teacher ratings of child behavior on the WPBIC. A sufficient observation period increases the reliability and validity of the teacher's ratings and also reduces the probability that such high magnitude, yet low frequency behaviors, as stealing, temper tantrums, and fighting will be missed by the teacher.

The scale can be most efficiently used if the teacher waits approximately two months after the start of school and then rates each child in her class. Children who are in need of specialized educational services or those who should be referred for further evaluation and treatment can be identified early in the school year. Additional behavioral problems which may develop in individual children as the year progresses can be rated on the WPBIC as they occur.

Development and Standardization of the WPBIC

Source of WPBIC Items: The fifty checklist items were drawn from teacher descriptions of classroom behavior problems. A random sample of thirty experienced teachers was drawn from the population of fourth, fifth, and sixth grade teachers in a local school district. The teachers were then asked to nominate those children in their classes who exhibited chronic behavior problems. Each teacher was then interviewed and asked to describe the child's behavior problem(s) and to give operational descriptions of the behaviors that concerned them. Observable descriptions of overt behavior were abstracted from each interview, yielding an item pool of three hundred items. Fifty of the most frequently mentioned behaviors from this sample were selected for inclusion in the scale.

Derivation of Item Score Weights: A panel of five behavioral scientists was selected and assigned an item rating task for the purpose of deriving score weights for individual scale items. The five judges were asked to rate each item's weight or influence in handicapping a given childs present adjustment. Judges rated each behavioral item's influence on a twenty point scale ranging from of no importance to great importance. The scale was a continuum on which the judges could rate an item at any given point. Judges' item ratings were pooled and averaged and each item assigned an arbitrary score weight ranging from four to one on the basis of such ratings. The results of the rating procedure and the assignment of score weights are presented in tables 1 and 2 below.

Table I

Mean Scores, Standard Deviations, and Inter-Rater Reliability (R_{11}) for All Judges on Fifty Scale Items

Judges	Mean	S.D.
#1	11.8	4.1
#2	9.5	3.6
#3	9.5	4.4
#4	11.6	3.7
#5	12.7	3.5
	Inter-judge reliability: .83	

Since r_{11} was .83, the means of the five judges on all items were pooled and assigned as score weights for the scale items.

Table 2

Item Mean Scores, Score Weights, and Number and Percentage of Items in Each Category

Mean Score	Score Wt.	N	%
16 15 14.4	4	6	12
13. 12.8	3	8	16
12 11.8	2	10	20
6.5	1	26	52
	Total	50	100

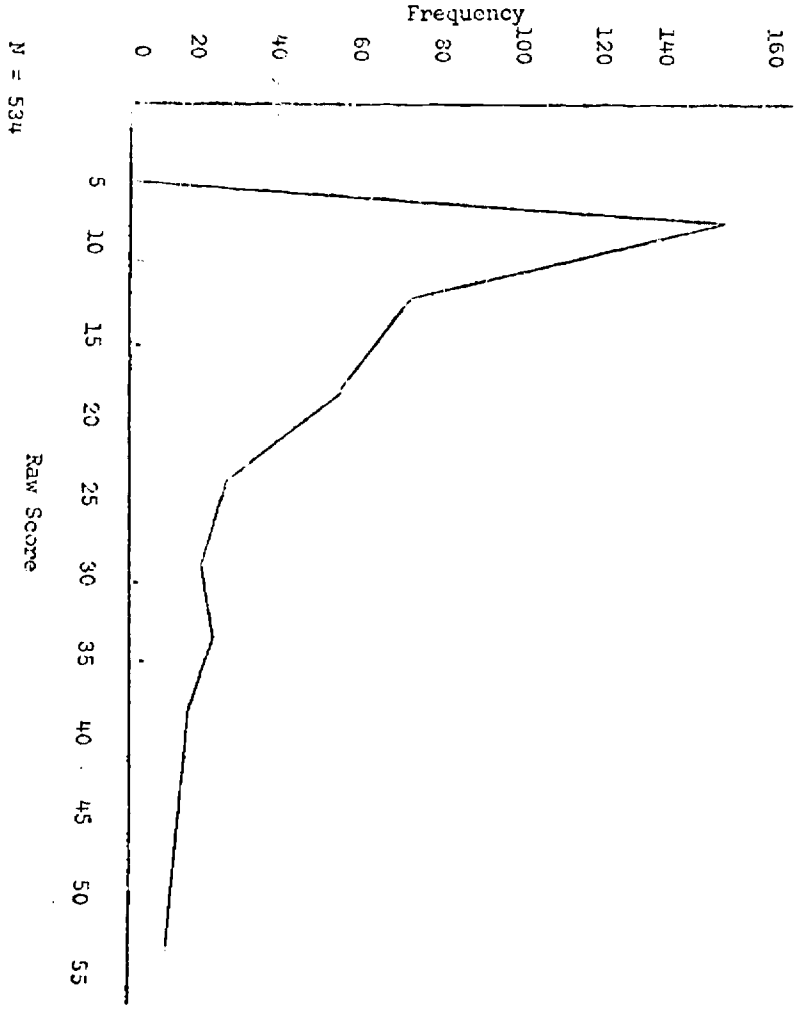
With this weighting system, a subject can receive a high score of one hundred and a low score of zero.

Normative Procedures: Items selected and weighted were incorporated into a behavior checklist and given to a twenty-one teacher sample of fourth, fifth, and sixth grade elementary teachers. Teachers evaluated all pupils in their classes on the checklist after having observed them for approximately two months in the classroom environment. Each subject evaluated on the scale received a marking of either present or absent for each item. Teachers were instructed not to single out problem children in their use of the scale since this would have undoubtedly biased results. This procedure yielded scores on 534 fourth, fifth, and sixth grade children. The mean score for the normative sample was 7.76 with a standard deviation of 10.53.

For purposes of screening and identification, it was necessary to select a point within the frequency distribution (checklist score) which would separate disturbed from nondisturbed children with an acceptable degree of reliability and validity. However, as noted in figure 1 below, the distribution of raw scores was positively skewed and did not represent a normal distribution.

Figure 1

Frequency Distribution of Raw Scores on Fifty Checklist Items



Since the WPBIC is composed of fifty negative behaviors, a positively skewed distribution would be expected when the scale is administered to a regular school population. However, in a residential treatment facility for severely disturbed children, the scale's application could conceivably result in a negatively skewed distribution as high scores indicate possession of a large number of deviant behaviors. Since behavioral adjustment is considered to be normally distributed in ordinary populations, the raw data on 534 subjects were converted into a T score distribution so as to normalize the data and to establish separation points within the distribution.

Table 3

Summary T-Score Conversion Table

T-Score	Raw Score
90	50
80	41
70	31
60	21
50	11
40	1

A T score of 60, which is the equivalent of one standard deviation above the mean, was established as the point in the distribution for separating disturbed from nondisturbed subjects. In using the WPBIC, subjects who receive a raw score of 21 (T score of 60) or above should be referred for a more intensive behavioral analysis and evaluation.

Reliability of the WPBIC

The reliability of the WPBIC was estimated by the Kuder-Richardson split-half method. The instrument was divided into equivalent split-halves by selecting odd and even numbered items for inclusion in the two half tests. In an effort to make the two halves of the scale more nearly equivalent and to reduce the response bias which operates when a group of deviant behaviors cluster together in serial form, items and their equivalent score weights were distributed equally among the two half tests. One behavior with a score weight of four was assigned as item number fifty and another behavior with a score weight of four was assigned as item number one. This procedure was duplicated for the remaining forty-eight items by alternately assigning score weights of four, three, two, and then one to the two halves of the scale. The split-half reliability coefficient obtained on the scale was .98 with a standard deviation of 10.53 and a standard error of measurement of 1.28. A coefficient of .98 indicates that 97% of the variance of test scores in the sample was true score variance and 3% is error variance. With a reliability coefficient of .98, the scale is capable of making individual comparisons among subjects with a considerable degree of reliability as an

r of .90 is the minimum coefficient acceptable for this purpose (Lindquist, 1951).

A formula was applied to the reliability coefficient in order to determine the effect upon reliability of the WPBIC by first doubling and then tripling its length. By this formula, a one hundred item scale would yield an r of .99 and a one hundred-fifty item scale would also yield an r of .99. The gain which would be realized by doubling or tripling the length of the present scale would be .01. The WPBIC, in its present form, appears to be at its near optimum length at fifty items.

Validity of the WPBIC

Four types of validity were estimated on the WPBIC: contrasted groups validity, criterion validity, factorial validity, and item validity. The validity data were derived from the original normative sample.

Contracted Groups Validity: In the contrasted groups method of assessing validity, two independent groups are defined in relation to the construct being measured and the instrument is then administered to both groups. Differences between the two groups in test score are then tested for statistical significance. (Levitt, 1960). Two independent groups were defined in relation to the construct of behavior disturbance. Thirty-eight subjects in the 534 pupil sample were identified as behaviorally disturbed according to one or more of the following criteria: 1. has been examined by a psychologist and referred to a psychiatric or clinical facility, 2. specific educational provisions have been made for the subject within the school setting because of his behavior problem(s), 3. has received instruction at home because of his inability to profit from classroom instruction due to his behavior problem(s). These thirty-eight subjects, so identified, were matched with thirty-eight subjects from the normative sample, not so identified, in terms of age, grade, and sex. All pupils who matched the experimental subjects in age, grade, and sex were lifted from the sample. A table of random numbers was used to facilitate the random selection of thirty-eight control subjects to be paired with the experimental subjects for purposes of experimental analysis.

Table 4

Means, Standard Deviations, and N's of Experimental and Control Groups with Test for Statistical Significance

Experimental (N=38)		Control (N=38)		D	CR
$\bar{X}$	S.D.	$\bar{X}$	S.D.		
16.03	12.68	6.47	5.47	10.16	4.23*

*Significant beyond .001 level

The difference between the means of the experimental and control subjects was significant beyond the .001 level of confidence. Contrasted groups validity can be reasonably claimed for the WPBIC since behaviorally disturbed subjects received significantly higher scores on the construct which the scale measures than did nonbehaviorally disturbed subjects.

Criterion Validity: A biserial correlation was computed on the normative data to assess the degree of relationship which exists between scores on the WPBIC and the construct of behavior disturbance as measured by the three criteria discussed above. If the scale measures disturbed behavior, then it appears reasonable to expect that scores of subjects who have been referred to psychiatric or clinical facilities or those who require special educational provisions because of such behavior problems should correlate higher with the criteria of behavior disturbance than scores of subjects who are judged not in need of such attention.

The biserial correlation between checklist score and the criterion yielded on r_{bis} of .68. The standard error of this correlation is .039 and its index of predictive efficiency is .33. The r_{bis} of .68 is significantly different from zero at the .01 level. The predictive efficiency index of .33 provides a measure of the scale's predictive value and indicates that the WPBIC has utility in the prediction of behavior disturbance in populations of elementary school children.

Factorial Validity: Data obtained from administering the WPBIC to a 534 pupil normative sample were factor analyzed according to a diagonalization method orginated by Jacobi and adapted by von Neumann for large computers, Ralston and Wilf (1962). The factors were then subjected to a Varimax Orthogonal rotation to obtain a simple structure. This procedure yielded five factors which are presented in Table 5 along with their constituent items and factor loadings.

Table 5

WPBIC Factors, Items, and Factor Loadings for a Sample of 534 Public School Pupils

Factor	Item	Factor Loading
1. Acting-out Syndrome disruptive, aggressive, defiant (14 items)	1	.53
	4	.74
	12	.56
	16	.55
	18	.72
	21	.49
	27	.63
	30	.63
	31	.60
	32	.69

	35	.64
	38	.39
	39	.74
	46	.77
2. Withdrawal Syndrome (restricted functioning, avoidance behavior) (5 items)	15	.54
	29	.57
	37	.67
	42	.75
	45	.79
3. Distractability (short attention span, inadequate study skills, non-attending) (11 items)	3	.49
	6	.30
	9	.81
	10	.49
	13	.69
	14	.40
	19	.41
	24	.35
	41	.67
	49	.79
	50	.46
4. Disturbed Peer Relations (Inadequate social skills, negative self-image, compulsive) (10 items)	5	.50
	7	.55
	23	.61
	25	.73
	26	.72
	28	.56
	34	.33
	40	.77
	43	.55
	48	.05
5. Immaturity (dependent) (10 items)	2	.32
	8	.56
	11	.59
	17	.69
	20	.67
	22	.79
	33	.73
	36	.74
	44	.35
	47	.82

The results of this analysis are similar to the factors obtained by Quay, Morse and Cutler (1966) on a sample of emotionally disturbed children in special classes and by Patterson (1964) on a sample of children referred to a child guidance clinic. This type of analysis is useful in establishing the validity of an instrument since it provides specific information about the content of a scale (what the scale measures) and also provides for a more detailed description of behavior through factorial, profile analysis techniques. (See administration and scoring).

The relationships which exist between the item clusters that make up the five factors of the WPBIC are presented in the correlation matrix below.

Table 6

Inter-correlations of
the Five WPBIC Factors

	Acting Out Syndrome	Withdrawal Syndrome	Distractability	Disturbed Peer Relations	Immaturity
Acting Out Syndrome	--	.02	.67	.48	.39
Withdrawal Syndrome			.12	.18	.23
Distractability				.48	.44
Disturbed Peer Relations					.34
Immaturity					--

The correlations indicate that with the exception of item clusters one and three, there is very little overlap among the five factors. The factors seem to be relatively independent of one another. This suggests that the WPBIC measures separate functions of the same behavior domain (e.g. behavior disturbance).

The r of .67 between acting out syndrome and distractability indicates that 44 percent of either factor is attributable to overlap or common factor variance. The content of the items in each factor supports the assumption that the two factors represent common elements. In addition, acting out or hyperactive children often manifest very high rates of non-attending and distractive behavior (Walker and Buckley, 1968; Patterson et. al, 1965).

The normative sample of 534 subjects was scored on the five factors in order tain normed scores for each factor. These data were then converted score distributions for each of the five factors.

Table 7

Distribution of Raw Scores, T Scores, and Cumulative Percentages on the WPBIC for 534 Subjects

	Acting Out Syn.		Withdrawal Syn.		Distract.		Dist. Peer Rel.		Immaturity	
T Score	Raw Score	Cum%	Raw Score	Cum%	Raw Sc.	Cum%	Raw Sc.	Cum%	Raw Sc.	Cum%
100									10	100
--	26	100								
--										
--										
--							11	100	9	99
95	24	99								
--										
--										
--	23	99						99		
--	22	99					10		8	99
90										
--	21	98								
--			14	100			9	98		
--	20	98								
--									7	97
85			13	99						
--	19	97								
--							8	98		
--	18	97								
--					13	100				97
80	17	96		98					6	
--			11				7	97		
--	16	96			12	99				
--										
--	15	95	10	97						
75					11	98				
--	14	94					6	97	5	96
--			9	94						
--	13	94			10	96				
70	12	94	8	93						
--					9	93	5	95	4	94
--	11	93								
--										
--	10	92	7	91	8	91				
65							4	94		
--	9	92								
--			6	98					3	92
--	8	91			7	87				
--										
60	7	89	5	87	6	84	3	90		
--										
--									2	88
	6	87		84	5	81				

Table 7 Continued

T Sc.	Acting Out Syn.		Withdrawal Syn		Distractabil.		Dist. Peer Re		Immaturity	
	Raw Sc.	Cum %	Raw Score	Cum %	Raw Sc.	Cum%	Raw Sc.	Cum%	Raw Sc	Cum %
--			4							
55	5	85					2	86		
--			3	81	4	76				
--										
--	4	83								
--									1	84
50	3	81	2	78	3	71	1	65		
--										
--	2	75	1	76	2	64				
--										83
--	1	74					0	80	0	
45	0	67		70	1	50				
--			0							
--										
--						42				
--					0					
40										
	$\bar{X}$=2.23		$\bar{X}$=1.60		$\bar{X}$=2.63		$\bar{X}$=.78		$\bar{X}$=.65	
	sd=4.79		sd=3.19		sd=3.31		sd=2.16		sd=1.74	

Item Validity: Item variance indices, item validity indices, and item intercorrelations were computed on all fifty items of the WPBIC. The maximum variance (.25) which an item can have is the point at which the item can make the greatest number of separations among individuals. Garnett (1962) recommends item variance values of .24-.25 for most educational test items since it is desirable to make maximum separations among individuals in terms of mental ability, aptitude, and achievement factors. However, when constructing an instrument which will separate a predetermined portion of individuals from the total sample, the .24-.25 value for optimal selection of items does not apply. With the WPBIC, it was important to select items which were not so narrow or limited in scope that they were useless for purposes of identification. On the other hand, a behavior such as not paying attention, is so common and so general that it is probably typical of most school children at one time or another. This behaviors i nocuous content and extremely high frequency would, in all likelihood, negate its value in the identification process. Since approximately ten to twenty percent of school children have serious behavior problems, a criterion for WPBIC item selection, or the basis of variance indices, was established at from .09 to .16. A value of .09 equals ten percent possessing a behavior and .16 is equal to twenty percent possessing the behavior.

Table 8

Item Variance and Standard Deviation Indices for Fifty Checklist Items

Item	Variance Index	S.D.
1	.12	.69
2	.05	.47
3	.15	.78
4	.08	.58
5	.01	.25
6	.09	.60
7	.02	.29
8	.04	.43
9	.21	.92
10	.14	.78
11	.01	.28
12	.05	.50
13	.17	.85
14	.14	.76
15	.13	.74
16	.05	.48
17	.02	.33
18	.09	.63
19	.11	.67
20	.04	.45
21	.03	.39
22	.01	.22
23	.12	.33
24	.12	.70
25	.02	.28
26	.02	.30
27	.04	.45
28	.03	.43
29	.09	.63
30	.04	.43
31	.03	.36
32	.05	.50
33	.00	.12
34	.01	.22
35	.12	.72
36	.00	.12
37	.06	.51
38	.13	.73
39	.07	.55
40	.05	.48
	.17	.84

Table 8 Continued

Item	Variance Index	S.D.
42	.08	.59
43	.04	.45
44	.01	.25
45	.12	.73
46	.04	.44
47	.00	.17
48	.03	.36
49	.21	.93
50	.10	.66

The range of item variance indices is from .00 to .21 and the item standard deviations range from .12 to .93. Seventeen of the items have variance indices which fall within the optimal range of .09 to .16 for the separation of the disturbed segment of the school population (approximately ten to twenty percent) from the remainder of the school population. The remaining variance indices fall either slightly below or slightly above this range with the exception of items 33, 36 and 47. The WPBIC items thus closely approximate the criterion of .09 to .16 chosen for judging the variance indices of individual items.

The intercorrelations among fifty scale items yeilded 1,225 coefficients which ranged in magnitude from .00 to .83. With the exception of several items the results of this analysis confirm the hypotheses that the WPBIC scale items are measuring separate functions of the same behavior domain and are not excessively duplicating one another's functions. This analysis also provides an empirical basis for evaluating the teacher's judgment of behavior problems in children. For instance, item #35 reads: "openly strikes back with angry behavior to teasing of other children" and item #42 reads: "doesn't protest when others hurt, tease, or criticize him." These two behaviors, by definition, would appear to be incompatible within the same subject. These two items intercorrelated at a value of -.03. Similarly, item #6 reads: "perfectionistic: meticulous about having everything exactly right" and item #7 reads: "will destroy or take apart something he has made rather than show it or ask to have it displayed." These two behaviors appear to be logically unrelated and the correlation between them should be low. Items #6 and #7 intercorrelate at a value of .00. This result is especially significant in view of the fact that adjacent items ordinarily intercorrelate highly as a function of response set. At the other extreme, items #9 and #49 both measure distractive behavior and intercorrelate at a value of .83. With this amount of duplication, either item could perform the function of the other.

A biserial correlation between scale items and the total score was computed yielding a discrimination index which is a measure of internal consistency

between individual items and test score. The specific procedure involved the selection of upper and lower groups, in terms of checklist score, according to Kelley's (1939) criteria for the validation of test items and then correlating each item with total score which served as the criterion variable.

Table 9

Item Validity Indices
on Fifty Checklist Items

Item	Validity Index
1	.67 * *
2	.19 * *
3	.67 * *
4	.65 * *
5	.33 * *
6	.09
7	.45 * *
8	.42 * *
9	.54 * *
10	.61 * *
11	.24 * *
12	.49 * *
13	.48 * *
14	.65 * *
15	.14 *
16	.55 * *
17	.19 * *
18	.59 * *
19	.52 * *
20	.33 * *
21	.48 * *
22	.12 *
23	.39 * *
24	.56 * *
25	.40 * *
26	.35 * *
27	.58 * *
28	.48 * *
29	.40 * *
30	.57 * *
31	.42 * *
32	.60 * *
33	.10
34	.26 * *
35	.62 * *
[illegible]	.10
[illegible]	.28 * *
[illegible]	.55 * *
[illegible]	.59 * *

Table 9 Continued

Item	Validity Index
40	.30 **
41	.53 **
42	.12 *
43	.39 **
44	.15 **
45	.36 **
46	.59 **
47	.03
48	.15 **
49	.58 **
50	.32 **

** Significant at .01 level * Significant at .05 level

The item validity indices on the fifty items vary from .03 to .67. The validity indices indicate that the individual items correlate highly with the criterion (total score) and that the items discriminate between subjects in the upper and lower twenty-seven percent of the sample in terms of checklist score. The item validities suggest further that the items making up the WPBIC constitute a very homogenous, related set of behaviors with the exception of items 33, 36 and 47 which have indices of .10, .10, and .03 respectively.

Educationally Related Variables

Hypotheses were constructed to determine the effect which non-behavioral but educationally relevant variables have upon WPBIC scores of subjects in the study sample. These variables include grade of student, sex of student, and sex of rater.

Table 10

Sex Differences in Checklist Score on all Subjects

Male (N = 276)		Female (N = 259)			
$\bar{X}$	S.D.	$\bar{X}$	S.D.	D	Critical Ratio
10.50	12.16	4.83	7.40	5.67	6.67 **

* Significant at .05 level ** Significant at .01 level

Table 11

Grade Differences in Checklist Score on all Subjects

(N = 164) Grade 4		(N = 196) Grade 5		(N = 174) Grade 6				
$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.	F Ratio	D	CR
						11.23**		
9.48	11.26	8.72	11.87				.76	.62
9.48	11.26			5.04	7.28		4.44	4.23**
		8.72	11.87	5.04	7.28		3.68	3.64**

* Significant at .05 level ** Significant at .01 level

Table 12

Score Differences by Sex of Rater on all Subjects

Male Rater (N = 10)		Female Rater (N = 10)			
$\bar{X}$	S.D.	$\bar{X}$	S.D.	D	CR
7.12	10.53	8.43	10.39	1.31	1.47

* Significant at .05 level ** Significant at .01 level

Table 13

Score Differences When Subjects Are Rated by a Rater of the Same Sex Versus a Rater of the Opposite Sex

Rating Comparisons	N	$\bar{X}$	S.D.	F Ratio	D	CR
Male (R) rates Male (S)	148	9.60	12.80	17.67**	1.97	1.85
Female (R) rates Male (S)	127	11.57	11.04			
Male (R) rates Female (S)	128	4.26	7.41		1.72	1.89
Female (R) rates Female (S)	129	5.98	7.00			
Male (R) rates Male (S)	148	9.60	12.80			
Female (R) rates Female (S)	129	5.98	7.00		4.62	3.81**
Male (R) rates Male (S)	148	9.60	7.41			
Male (R) rates Female (S)	128	4.26	7.41		5.34	----**
(R) rates Female (S)	129	5.98	7.00			
(R) rates Male (S)	127	11.57	11.04		5.59	----**

Table 13 Continued

Rating Comparisons	N	X̄	S.D.	F Ratio	D	CR
Female (R) rates Male (S)	127	11.57	11.04			
Male (R) rates Female (S)	128	4.26	7.41		7.31	----**

* Significant at .05 level ** Significant at .01 level

Table 13

Sex Differences on all Subjects by Grade

Grade of S	Male			Female			F Ratio	D	CR
	N	X̄	S.D.	N	X̄	S.D.			
							14.25**		
Grade 4	87	12.02	13.63	77	6.62	9.00		5.40	3.13**
Grade 5	102	12.63	14.03	94	4.47	6.92		8.16	----**
Grade 6	86	6.54	7.81	87	3.62	5.74		2.92	2.87**

* Significant at .05 level ** Significant at .01 level

In Table 10, it can be seen that male students received significantly higher scores on the WPBIC than female students. This result is consistent with research findings which have indicated that significantly higher proportions of boys than girls are identified as behaviorally disturbed. Beilin (1959). This finding also strengthens the applicability of the scale for use with school populations in that the checklist reflects sex differences in behavior disturbance which are known to exist in such populations.

In Table 11, the analysis indicates that sixth grade students were rated as significantly less deviant than either fifth or fourth grade students. There is no empirical evidence, of which the writer is aware, that supports this finding. The result may be explained by the fact that the difference obtained represents a type one error in that no actual differences exist between the two groups even though the data appears to support the opposite conclusion. If this explanation were correct, then the null hypothesis would have to be accepted instead of rejected for this mean difference. Since the critical ratios between both fourth and sixth and fifth and sixth grade subjects were significant beyond the .01 level, this explanation is possible but highly improbable. Another explanation may be that sixth grade students are rated as less deviant than fourth and fifth grade students because of some as yet unexplained and unresearched maturational processes. A third possible explanation may be that the teachers who rated sixth grade students in this study were "easier" raters than fourth and fifth grade teachers.

No statistically significant differences were found between male and female raters on their ratings of all subjects. This result indicates, as would be expected, that male raters did not rate subjects as significantly more or

less deviant than female raters.

An analysis of variance applied to the means of subjects rated by a rater of the same sex and subjects rated by a rater of the opposite sex yielded an F ratio which was significant beyond the .01 level. However, inspection of the respective means indicates that male and female raters do not rate male subjects in a significantly different fashion; nor do male and female raters rate female subjects in a significantly different fashion. Thus, a same sex bias did not appear to be operate in the ratings of teachers in this sample. The major part of the variance is accounted for by the fact that both male and female teachers rated male students as significantly more deviant than female students.

The analysis in Table 13 for sex differences across grades four, five and six yielded an F ratio which is significant beyond the .01 level. Inspection of the means reveals that sex differences between male and female subjects in terms of checklist score, held constant across the three grades. It should be noted that even though sixth grade subjects were rated as significantly less deviant than fourth and fifth grade subjects, sex differences between male and female subjects in grade six were statistically significant.

Administration and Scoring

Scoring: The WPBIC is scored by counting the number of items marked present in each of the four columns on the form. Items marked present in each of the columns are then multiplied times their score weights and the totals entered in the appropriate squares at the end of the form. For example, the number of present items in column one is multiplied times one and this figure is entered in the weighted score box. The number of present items in column two is multiplied times two and this figure is entered in the weighted score box under column two. This same procedure is repeated for columns three and four. The sub-totals in the four boxes are then added and entered in the total weighted score box. This figure is the total score for a subject on the WPBIC.

Profile Analysis Chart: If a subject receives a weighted score of 21 (T-Score of 60) or above, then he is classified as disturbed and the profile analysis chart should be completed on his WPBIC ratings below. If the subject receives a weighted score of less than 21, then he is not classified as disturbed and the profile should not be completed on his WPBIC ratings.

The profile analysis chart (PAC) should be completed on all subjects who are referred for further analysis, evaluation, or treatment. The PAC provides the receiving agency or professional with specific information about the child's behavior disorder. For example, one child may be high in acting-out syndrome and distractability and low in the other factors while another may be high in withdrawal syndrome and low in the other four factors.

A subject who receives a raw score equivalent to a T score of 60 on any of the factors is considered to be high in the behavioral area defined by the items in that factor. Thus a child who is high in acting out syndrome would require a different intervention program than one who is high in withdrawal [illegible]. The PAC is intended to facilitate decision making in the areas of

psychological/educational diagnosis, evaluation and treatment.

References

1.Stouffer, G. A. Behavior problems of children as viewed by teachers and mental hygienists: A study of present attitudes as compared with those reported by E. K. Wickman. Mental Hygiene. Vol. 36, 1952, 271-285.

2.Beilin, H. Teachers and clinicians' attitudes toward the behavior problems of children--a reappraisal. Child Development, Vol. 30, 9-25.

3.Bower, E. A process for early identification of emotionally disturbed children. Bulletin of the California State Department of Education, Vol. 27 1958, 1-111.

4.Haes, W. R. The identification of emotionally disturbed children. Vol. 32, May, 1966, 607-613.

5.Quay, H. C., Morse, W. C., and Cutler, R. L. Personality patterns of pupils in special classes for the emotionally disturbed. Exceptional Children, 1966, 32, 297-301.

6.Patterson, G. R. An empirical approach to the classification of disturbed children. Journal of Clinical Psychology, 1964, 20, 326-337.

7.Patterson, G. R., Jones, R., Wright, M. and Whittier, J. A behavior modification technique for the hyperactive child. Behavior Research and Therapy, 1965, 2, 217-226.

8.Garrett, H. E. Statistics in Psychology and Education, New York: David McKay Co., Inc., 1962.

9.Kelley, T. L. The selection of upper and lower groups for the validation of test items. Journal of Educational Psychology, Vol. 30, 1939, 17-24.

10.Ralston, A. and Wilf, H. (Ed.) Mathematical Methods for Digital Computers John Wiley and Sons, New York, 1962, Chap. 7.

11.Walker, H. M. and Buckley, N. K. The use of positive reinforcement in conditioning attending behavior. Journal of Applied Behavior Analysis. Vol. 1 # 3, 245-250.

WPBIC Profile Analysis Chart

Factor I: Acting out Syndrome

Items (P)		Score Weight	Weighted Score
1.____	X	(3)	_____
4.____	X	(3)	_____
12.____	X	(1)	_____
16.____	X	(2)	_____
18.____	X	(1)	_____
21.____	X	(1)	_____
27.____	X	(2)	_____
30.____	X	(2)	_____
31.____	X	(4)	_____
32.____	X	(1)	_____
35.____	X	(3)	_____
38.____	X	(1)	_____
39.____	X	(1)	_____
46.____	X	(1)	_____
(Total) Weighted Factor Score			_____

Factor II: Withdrawal Syndrome

Items (P)		Score Weight	Weighted Score
15.____	X	(1)	_____
29.____	X	(2)	_____
37.____	X	(4)	_____
42.____	X	(3)	_____
45.____	X	(4)	_____
(Total) Weighted Factor Score			_____

Factor III: Distractability

Items (P)		Score Weight	Weighted Score
3.____	X	(1)	_____
6.____	X	(3)	_____
9.____	X	(4)	_____
10.____	X	(2)	_____
13.____	X	(4)	_____
14.____	X	(3)	_____
19.____	X	(3)	_____
24.____	X	(1)	_____
41.____	X	(3)	_____
49.____	X	(1)	_____

(Total) Weighted Factor Score _____

Factor IV: Disturbed Peer Relations

Items (P)		Score Weight	Weighted Score
5.____	X	(1)	_____
7.____	X	(3)	_____
23.____	X	(4)	_____
25.____	X	(2)	_____
26.____	X	(4)	_____
28.____	X	(3)	_____
34.____	X	(3)	_____
40.____	X	(1)	_____
43.____	X	(3)	_____
48.____	X	(1)	_____

(Total) Weighted Factor Score _____

Factor V: Immaturity

Items (P)		Score Weight	Weighted Score
2.____	X	(2)	_____
8.____	X	(4)	_____
11.____	X	(2)	_____
____	X	(3)	_____
____	X	(3)	_____

Items		Score Weight	Weighted Score
(P)			
22.____	X	(1)	____
33.____	X	(1)	____
36.____	X	(1)	____
44.____	X	(1)	____
47.____	X	(1)	____

(Total) Weighted Factor Score____

Plot the weighted factor scores on the chart below to form a profile analysis and connect them with a straight line. A horizontal line has been drawn across the chart corresponding to a T Score of 60(one s.d. above the mean score of the norm sample for each factor.) If a subject receives a raw score which corresponds to a T score of 60 or above, he is considered to be high on that factor(s) and could benefit from a treatment program designed to remediate behavior disorders represented by that factor.

Profile Analysis

T-Score	Acting out Syn. Raw Score	Withdrawal Syn. Raw Score	Distract. Raw Sco.	Dist. Peer Rel. Raw Score	Immaturity Raw Score
100					
--	26				
--					
--				11	
--					
95	24				9
--					
--					
--	23				
--					
90	22			10	8
--	21	14			
--					
--	20			9	
--					
85		13			7
--	19				
--				8	
--	18		13		
--					
80	17	11			
--					
--	16		12	7	6
--					
--	15	10			
75			11		
-	14			6	5
--					
--	13	9	10		
--					
70	12	8			
--			9	5	4
--	11				
--					
--	10	7	8		
65					
--	9				
--		6	7	4	3
--					
--	8				
60		5	6	3	
--	7				
--			5		2
--	6	4			
--				2	

	Acting out Syn.	Withdrawal Syn.	Distract.	Dist. Peer Rel.	Immat.
T-Score	Raw Score	Raw Score	Raw Score	Raw Score	R. Sc.
55	5		4		
--		3			
--					
--	4		3		
--	3				1
50		2		1	
--	2		2		
		1			
--	1				
--				0	0
45	0		1		
--					
--		0			
--					
--			0		
40					

The Walker Problem Behavior Identification Checklist

by

Hill M. Walker, Ph.D.

NAME	SCHOOL	
ADDRESS	GRADE	
BIRTHDATE	AGE	RATER
DATE	MALE	FEMALE

General Instructions

Please read each item carefully and respond by placing a check (✓) in the present or absent column as it applies to the child. If you have observed a behavioral item in the child's response pattern during the last two month period, answer the item by marking in the present column. If you have not observed the behavior in the child during this period, mark in the absent column. Mark either present or absent for each item. Do not omit any.

Examples:

	Sc. Wt. 1		Sc. Wt. 2		Sc. Wt. 3		Sc. Wt. 4	
	P	A	P	A	P	A	P	A
1. Has temper tantrums	✓							
2. Has no friends							✓	
3. Refers to himself as dumb, stupid or incapable					✓			
4. Must have approval for tasks attempted or completed	✓							

Items #1 and #4 are rated as present while items #2 and #3 are rated as absent. Items marked as present are scored and multiplied times their score weights. Items marked absent are not scored. (See administration and scoring for directions).

Items	Sc. Wt. 1		Sc. Wt. 2		Sc. Wt. 3		Sc. Wt. 4	
	P	A	P	A	P	A	P	A
1. Complains about others' unfairness and/or discrimination towards him.								
2. Is listless and continually tired.								
3. Does not conform to limits on his own without control from others.								
4. Becomes hysterical, upset or angry when things do not go his way.								
5. Comments that no one understands him.								
6. Perfectionistic: Meticulous about having everything exactly right.								
7. Will destroy or take apart something he has made rather than show it or ask to have it displayed.								
8. Other children act as if he were taboo or tainted.								
9. Has difficulty concentrating for any length of time.								
10. Is overactive, restless, and/or continually shifting body positions.								
11. Apologizes repeatedly for himself and/or his behavior.								
12. Distorts the truth by making statements contrary to fact.								
13. Underachieving: Performs below his demonstrated ability level.								
Disturbs other children: teasing provoking fights, interrupting others.								

	Sc. Wt. 1		Sc. Wt. 2		Sc Wt. 3		Sc. Wt. 4	
	P	A	P	A	P	A	P	A
15. Tries to avoid calling attention to himself.								
16. Makes distrustful or suspicious remarks about actions of others toward him.								
17. Reacts to stressful situations or changes in routine with general body aches, head or stomach aches; nausea.								
18. Argues and must have the last word in verbal exchange.								
19. Approaches new tasks and situations with an "I can't do it" response.								
20. Has nervous tics: muscle-twitching, eye-blinking, nail-biting, hand-wringing.								
21. Habitually rejects the school experience through actions or comments.								
22. Has enuresis.								
23. Utters nonsense syllables and/or babbles to himself.								
24. Continually seeks attention.								
25. Comments that nobody likes him.								
26. Repeats one idea, thought, or activity over and over.								
27. Has temper tantrums.								
28. Refers to himself as dumb, stupid, or incapable.								
29. Does not engage in group activities.								

	Sc. Wt. 1		Sc. Wt. 2		Sc. Wt. 3		Sc. Wt. 4	
	P	A	P	A	P	A	P	A
30. When teased or irritated by other children, takes out his frustration(s) on another inappropriate person or thing.								
31. Has rapid mood shifts: depressed one moment, manic the next.								
32. Does not obey until threatened with punishment.								
33. Complains of nightmares, bad dreams.								
34. Expresses concern about being lonely, unhappy.								
35. Openly strikes back with angry behavior to teasing of other children.								
36. Expresses concern about something terrible or horrible happening to him.								
37. Has no friends.								
38. Must have approval for tasks attempted or completed.								
39. Displays physical aggression toward objects or persons.								
40. Is hypercritical of himself.								
41. Does not complete tasks attempted.								
42. Doesn't protest when others hurt, tease, or criticize him.								
43. Shuns or avoids hetero-sexual activities.								
44. Steals things from other children.								

Sc. Wt. 1		Sc. Wt. 2		Sc. Wt. 3		Sc. Wt. 4	
P	A	P	A	P	A	P	A

45. Does not initiate relationships with other children. ☐☐

46. Reacts with defiance to instructions or commands. ☐☐

47. Weeps or cries without provocation. ☐☐

48. Stutters, stammers, or blocks on saying words. ☐☐

49. Easily distracted away from the task at hand by ordinary classroom stimuli, i.e. minor movements of others, noises, etc. ☐☐

50. Frequently stares blankly into space and is unaware of his surroundings when doing so. ☐☐

	Sc. Wt. 1		Sc. Wt. 2		Sc. Wt. 3		Sc. Wt. 4	
	P	A	P	A	P	A	P	A
Number Items Marked Present								
Weighted Score								

Total Weighted Score ☐ + ☐ + ☐ + ☐ = ☐

Validation of a Behavior Rating Scale for Measuring Deviant Behavior Within the Classroom Setting[1]

Submitted for review and publication to
Journal of Consulting Psychology

Hill M. Walker
Department of Special Education
University of Oregon
Eugene, Oregon 97403

The classroom teacher usually experiences little difficulty in pinpointing those pupils whose behavior is deviant or disruptive to the classroom setting. The acting-out child is generally obvious to the casual observer of the classroom. There are, however, children in the classroom setting with equally severe behavioral handicaps who are not so easily identified. Children with behavioral deficits resulting in social withdrawal are not so obvious to the teacher or casual observer. However, their need for remediation and special services is just as great as the acting out child's.

A standardized method of observation and assessment in which equal attention of the observer or rater is focused on acting-out as well as social withdrawal behaviors would seem desirable. There are a number of scales available for rating behavior such as those described by Ross, Lacey, and Parton (1965), Becker, 1960, Novick, Rosenfeld, Black, and Dawson (1966), Dreger, (1964), Quay and Quay (1965) and Cromwell and Davis (1965). Many of these scales appear to have a high degree of face validity. However, there are important limitations associated with each. For example, one scale is designed for use with boys only; another is excessively long; and a third is validated with parents as raters. All these scales appear to be loaded toward identifying the child who exhibits deviant disruptive behavior.

The author has designed a multi-dimensional assessment model of increasingly refined levels of observation and assessment for identifying deviant behavior in children (Walker, 1969). The instruments are: a 50 item behavior problem checklist (WPBIC) which is used as an initial screening device (Walker, 1970); a behavior rating scale (BRS) of 62

item length for collecting data on child behavior as well as the teacher's response and reaction to that behavior; and a behavioral observation form for recording task-oriented behavior. The model is designed to give equal attention to the identification and description of acting-out and social withdrawal behavior in children. The development of these instruments and the validation of the checklist was described in an earlier paper (Walker, 1969). This study describes the validation of the behavior rating scale (stage two of the assessment model) as an instrument to identify deviant behavior in children.

Method

Subjects

The scales used in the validation procedures were administered by 35 teachers in grades one through six from three elementary schools of a local district. One of the teachers taught a class of ten year old educable mentally retarded children. Each rater was paid for participating in the study upon satisfactory completion of four sets of ratings on all children in his class. Teachers in two of the schools heard of the program through the principal at a staff meeting. The other teachers were personally asked to participate by their principal. A total of 356 children were rated in the study.

Tests

The test being validated, the behavior rating scale (BRS), is a 62 item scale containing descriptions of behavior that can be observed in the classroom setting. The first part of the scale consists of 32 [illegible] each requiring three rating judgments: rate of occurrence of

the behavior, rater response to the behavior, and behavioral effects or rater reaction to the behavior. The second part of the scale requires rating judgments only on rate of occurrence of deviant behavior. All judgments are made on a five point scale from 0-5. Two equivalent forms of this scale were constructed. Each scale consists of 62 items.

A pool of 189 items describing classroom behavior were used to construct the two equivalent forms of the BRS. A panel of behavioral scientists, composed of a school psychologist, a remedial teacher, a social worker, a psychologist, and a child psychiatrist was asked to sort the 189 behaviors into educationally relevant behavioral categories. The expected outcome of the sorting task was a behavioral classification of the scale items that would be educationally prescriptive and that would facilitate treatment decisions and referrals by psychological personnel in the school setting. After construction of the behavior classification system, these items or behavioral statements were further refined and incorporated into a behavior rating scale designed to provide data on associated dimensions of deviant behavior.

Three response measures are obtained on each item in the scale. Rate of occurrence provides a measure of the frequency with which a given behavior occurs over time. Rater response indicates how the teacher (or rater) responds to different behaviors as they occur within the educational setting. Rater reaction indicates the extent to which a given teacher is disturbed or irritated by deviant behaviors produced within the classroom setting. The rationale for this response measure grows out of the hypothesis that deviant behaviors that are highly irritating or disturbing to the classroom teacher are significantly

more predictive of an educational or psychological referral than are equally handicapping deviant behaviors which are less disturbing for the teacher. The rating criteria and a sample item are presented below:

Indicate your judgments in each of the three scoring areas according to the following criteria.

Section A: Rate of Occurrence

(0) The behavior has never occurred.
(1) The behavior occurs at least once every two months.
(2) The behavior occurs at least once a month.
(3) The behavior occurs at least once a week.
(4) The behavior occurs at least once a day.
(5) The behavior occurs at a constant or near constant rate.

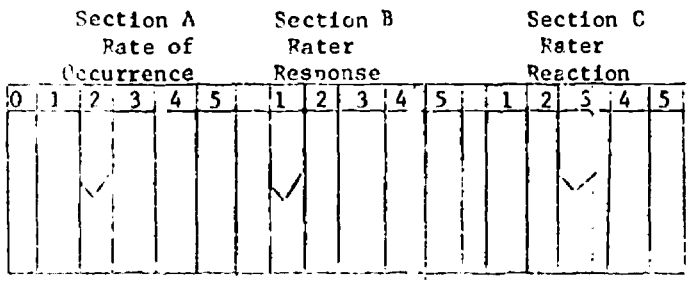
Section B. Rater Response: When this particular behavior occurs, do you

(1) Ignore the behavior?
(2) Give the child a warning glance?
(3) Interact verbally or physically with the child?
(4) Temporarily remove the child from the classroom setting?
(5) Refer the child to an outside source, i.e., counselor, psychologist, or separate referral agency?

Section C: Rater Reaction

(1) The behavior does not disturb you.
(2) The behavior disturbs you to a slight extent.
(3) The behavior disturbs you to a moderate extent.
(4) The behavior disturbs you to a great extent.
(5) The behavior disturbs you to a very great extent.

Sample Item:



This behavior is rated as: occurring at least once a month; the rater ignores the behavior; the behavior is moderately disturbing to the rater.

The Devereaux scale for disturbed children (DESB) was correlated with the behavior rating scale to obtain a measure of the concurrent validity of the BRS. This test, like the BRS, was designed for use by teachers in assessing deviant behavior in children within the educational setting. The Devereaux scale is composed of 47 items, 26 of which have a five point rating scale (1-5) and 21 which have a seven point rating scale (1-7). The DESB has a test-retest reliability for factors of .87 with the standard errors for factors ranging from 3.1-1.5. The test is supported by factor and contrasted group validity. The 47 items, each of which contributed most to one of twelve factors, were selected from a larger sample of items. Children with disturbed behavior (from the Devereaux School) were rated as significantly different from normal children on eleven of the twelve factors.

Concurrent validity was also assessed in terms of the BRS's relationship to the Walker Problem Behavior Identification Checklist (WPBIC). Since these two instruments are part of multi-stage assessment model, they should measure the same behavioral dimensions. The WPBIC consists of 50 descriptions of overt, deviant behavior which the teacher rates as being present or absent for each pupil. Items were assigned one of four score weights, from 1 to 4, indicating to what extent possession of a behavioral item handicaps adjustment. The Kuder-Richardson reliability of the WPBIC is .98. The average item validity is .40. Contrasted groups validity indicates there was a statistically significant difference between the mean score of a group of disturbed children and the mean score of a group of normal children. A biserial correlation between checklist scores and criterion scores

(based on three criteria of behavior disturbance) was .68.

The 15 teachers were divided into two groups by school. Six of the teachers rated their children on the BRS (forms I and II) and the Devereaux Scale. The nine remaining teachers rated children on the BRS (form I) and the WPBIC. The rating scales were left with the principal of each school along with a brief instruction sheet for each teacher. The teachers were to rate each child in their class on each of the two scales. All teachers were asked to rate all the children on the BRS before they rated any child on the comparable scale. Ten days after they were delivered, the completed forms were picked up. Three weeks later another set of rating scales was delivered. Group I teachers were given form II of the BRS (as well as the Devereaux Scale) and group II teachers were again given form II of the BRS and the WPBIC. Each teacher also received a letter commending the previous effort and reminding teachers of the necessity for identifying each child in the same way as before. The completed scales were collected after nine days.

Insert Table 1 About Here

From Table 1, it can be seen that this study was designed to enable calculation of test-retest reliabilities for the BRS (form II) the WPBIC, and the Devereaux Scale; equivalent forms reliability for BRS; and concurrent validity of the BRS using the Devereaux and WPBIC as criterion scales.

Results

Insert Table 2 About Here

The data in Table 2 indicate that means and standard deviations for both groups were smaller for rating time two than rating time one. This decrease is particularly evident for the means of the BRS which decreased from 16.9 to 11.2 in group 1 and from 41.3 to 29.4 in group 2. There was also a large difference in means and standard deviations of the BRS for groups 1 and 2 e.g. means of 16.9 and 11.2 for group 1 compared to means of 41.3 and 29.4 for group 2.

Two coefficients of reliability were calculated, the Pearson product moment (r) and the intra-class correlation (R).

Insert Table 3 About Here

The test-retest reliability coefficients varied widely among individual teachers. For example, for the BRS, R's ranged from .41-.88. For the WPBIC, the R's ranged from .43-.96. And for the Devereaux R's ranged from .46-.95. All the reliability coefficients shown in Table 3 are below the .90 figure suggested as the minimum degree of reliability desirable for selecting individuals from a group (Thorndike, 1951). If the tests were doubled in length, by the inclusion of valid items, the reliabilities would increase: BRS from .74 to .86, WPBIC from .80 to .89; Devereaux from .83 to .91. If the scales were tripled in length, BRS would have a reliability coefficient of .89, WPBIC .92, and Devereaux .94. Thus, if the scales were lengthened, they could a[illegible]e level of reliability suggested as desirable for this purpose b[illegible]like (1951).

As with the reliability coefficients, the validity coefficients varied among individual teachers.

Insert Table 4 About Here

For example, BRS (form I) correlations with the WPBIC ranged from .57-.90. In Table 4, the correlation of the BRS with the Devereaux was highest for form I (.81). The relationship of BRS (form II) with the WPBIC remained constant over time. As mentioned earlier, the criterion scales did not have the reliability deemed optimal by Thorndike (1951) for selecting individuals from a group. If the WPBIC and Devereaux scales had reliabilities of at least .90 then the BRS would correlate at least .84 and .75 with the Devereaux and .82 and .83 with the WPBIC.

Discussion

In this study, teachers were instructed to rate all pupils on the BRS before rating any child on the criterion scale in order to restrict the influence of the criterion test on the predictor test. However, the data in Table 2 would seem to indicate this instruction was not followed. The scoring system for the Devereaux scale resulted in much larger scores for deviant behavior than did the scoring procedure for the WPBIC. Those teachers who used the Devereaux rated their children much higher on the BRS than did teachers using the WPBIC as a comparison scale. There could possibly have been a set established for rating high or low on the BRS based on the scores of the criterion scale. As a matter of fact, many of the completed scales were grouped by child rather than by scale. Another explanation could be that the group using the Devereaux was more careful in its ratings than group 2.

(Group 1 was the group formed by the principal's personal request to each teacher). Finally, the difference in means could indicate that children in group 1's school produced higher rates of deviant behavior than children in group 2's school. Whatever the source(s) of these mean differences, the result indicates that care must be taken in establishing the criterion for identifying a child as deviant. It must be determined that it is the child's behavior and not the circumstances and setting in which he is rated which indicates that he is deviant. The three stage identification process of which the BRS is stage two, was designed to eliminate just this kind of uncertainty.

The differences in the means and variances over time could be due to a number of factors. The study was conducted during the last month of the school year. It is possible that the children's behavior improved in an attempt to increase their final school evaluations. The lower mean scores of deviant behavior on the second rating could also be due to decreased interest on the part of teachers. With many end-of-the-year reports to complete, it would be understandable if the teachers rated the children with less care on these scales the second time. However, this difference in mean score was larger for the BRS than for the criterion scales and could indicate a greater annoyance with the former. Not only does the BRS have more items than the other two scales, but half of those items require three different rating judgments. Is conceivable that if the BRS took longer to complete, it might be used with less care the second time than the criterion scales which are shorter and easier to use.

The differences in means and variance may appear to be large. However, the fact that the product moment and intraclass correlation coefficients of reliability are so similar would indicate that the differences were not crucial from time to time. The intraclass correlation coefficient indicates the degree of similarity of two observations rather than prediction of one observation from another. In the four parameter case used in this study, the R coefficient is based on means of the two tests, the product moment correlation between them, and a variance (the two tests are assumed to have the same variance). This paradigm would test the similarity of two observations, requiring that their variances be equal. In this study, the intraclass correlation would indicate that the BRS (form II), given at different times, and the two forms of the BRS given at different times, have a high degree of similarity.

It is interesting that the test-retest reliability of the BRS was lower than the equivalent forms reliability. In general, it is expected that equivalent forms reliability coefficients will be lower than test-retest reliabiltiy. There are two possible explanations for this result. Group 2, from whom data was collected for calculating the test-retest reliability, could have been a less conscientious group. Perhaps because they were personally asked to participate by the principal, group 1 teachers were more careful raters. Another explanation could be that repetition of the same scale was tedious and thus ratings were performed haphazardly. However, the fact that both the criterion scales had similarly high correlatives would tend to discredit these two

 ations.

The equivalent forms reliability coefficient of the BRS is high and the test-retest coefficient is adequate. However, neither reaches the high level of reliability suggested by Thorndike (1951) as being desirable for evaluating differences in a group (.90). The Spearman-Brown prophecy formula indicates theoretically that by doubling the length of the BRS, the equivalent forms reliability could be increased to the desirable level. However, the BRS is a longer form than either of the two criterion scales and it is also more difficult to use. Lengthening the test could, in the author's opinion, decrease its reliability because of the possibility of a fatigue factor.

The high validity coefficients would suggest that the BRS measures the same traits as the WPBIC and the Devereaux Scale. All the scales are aimed at measuring deviant behavior, but their approaches differ. The WPBIC contains very general items such as "underachieving; performs below his demonstrated ability level" or "has temper tantrums". The BRS expands on these items and gets at specifics such as "is easily thrown off and makes errors" or "when angry, slams books on the desk, etc." These two scales contain items which emphasize overt behaviors that the teacher can observe. The Devereaux Scale has items that are specific such as those in the BRS. However, the emphasis in this scale is on behavior that interferes or competes with academic performance such as: "rushes through work and therefore makes unnecessary mistakes," "gets openly disturbed about scores on a test." In general, items in all three scales appear to sample the same behavioral dimensions.

The BRS scale was designed to aid in the identification of children with deviant behavior. The BRS correlates highly with two other scales with known contrasted groups validity. The scale also records changes in deviant behavior rates as a result of treatment (Walker, 1969). However, additional research is needed to determine if the BRS discriminates between groups of normal and disturbed children.

References

Becker, W. C. The relationship of factors in parental ratings of self and each other to the behavior of kindergarten children as rated by mothers, fathers, and teachers. Journal of Consulting Psychology, 1960, 24, 507-527.

Cromwell, R. and Davis, D. Behavior classification of emotionally disturbed children. Paper presented at the meeting of the Council for Exceptional Children, Portland, Oregon, April 23, 1965.

Dreger, R. Behavioral classification project. Journal of Consulting Psychology, 1964, 28, 1-13.

Haggard, E. Intraclass Correlation and the Analysis of Variance, New York: Dryden Press, 1958.

Novick, J. Rosenfeld, E., Black, D., and Dawson, D. Ascertaining deviant behavior in children. Journal of Consulting Psychology 1966, 30, 230-238.

Quay, H. C. and Quay, L. C. Behavior problems in early adolescence. Child Development, 1965, 36, 215-220.

Ross, A. O., Lacey, H., and Parton, D. The development of a behavior checklist for boys. Child Development, 1965, 36, 1013-1027.

Spivack, G. and Surft, M. Patterns of disturbed classroom behavior May 1967, Final Report of Project No. 2156, Grant No. 32-48-7680-5023.

Thorndike, E. L. Reliability in E. Lindquist (Ed.) Educational Measurement, Washington D.C. American Council on Education, 1951, 560-621.

Walker, H. M. Empirical assessment of deviant behavior in children. Psychology in the Schools, 1969, 6, 93-97.

Walker, H. M. Walker Problem Behavior Identification Checklist. Western Psychological Services, 12031 Wilshire Blvd. Los Angeles, California.

Footnotes

1. This research was supported by U.S.O.E. Grant OEG 4-6-061308-0571 Assessment and Treatment of Deviant Behavior in Children.

2. Copies of BRS scales (forms I and II) can be obtained upon request from the author.

Table 1

Design of the Validation Study

	Rating Time 1 April 29-May 8	Rating Time 2 May 29-June 8
Group 1 6 Teachers 153 Children	BRS Form I Devereaux Scale	BRS Form II Devereaux Scale
Group 2 9 Teachers 206 Children	BRS Form II WPBIC	BRS Form II WPBIC

Table 2

Total Means and Standard Deviations
for Rating Times 1 and 2

	Group 1		Group 2	
	BRS	Devereaux	BRS	WPBIC
Means				
Rating Time 1	41.3(Form I)	121.6	16.9(Form I)	7.6
Rating Time 2	29.4(Form II)	119.7	11.2(Form II)	6.4
Standard Deviation				
Rating Time 1	30.9(Form I)	29.9	14.8(Form II)	9.3
Rating Time 2	27.6(Form II)	27.3	11.1(Form II)	7.9
N Time 1	153	153	200*	200*
N Time 2	153	153	206	206

*one teacher left 6 scales incomplete on time 1.

Table 3

Comparison of Interclass and
Intraclass Correlations Across Ratings

	Devereaux v Devereaux	BRS Form I v Form II	BRS Form I v Form II	WPBIC v WPBIC
r	.83	.83	.74	.80
R	.82	.83	.73	.78
N	153	153	200	200

Table 4

Concurrent Validity Coefficients

	BRS v Devereaux	BRS v WFBIC
r time 1	.81(Form I)(.84)*	.77(Form II) (.82)
r time 2	.72(Form II) (.75)	.78(Form II) (.83)

*corrected for attentuation

Stage Two

Behavior Rating Scale (BRS)

Scale 1

BEHAVIOR RATING SCALE

Demographic Information:

Name of Pupil_____ Date of Birth_____

School_____ Grade_____

Sex of Rater_____ Sex of Pupil_____

Name of Rater_____ Date_____

Instructions to the rater:

1 This scale is designed for the purpose of identifying behaviorally disturbed children. Items in the scale represent OVERT BEHAVIORS WHICH CAN BE VERIFIED BY OBSERVATION. Thus, if you have not observed a particular behavioral item in the classroom, you would indicate in the scoring section that the behavior had never occurred.

2. In the first part of the scale, three rating judgments are required for each behavioral item: (a) rate of occurrence, (b) rater response, (c) behavioral effects. One judgment is required under (a) rate of occurrence; one judgment is required under (b) rater response; and one judgment is required under (c) behavioral effects. Thus, there would not be more than three rating judgments per item.

Rate of occurrence is designed to secure information on the frequency with which a particular behavior occurs within the classroom setting. For example, if a behavior occurs one or more times in a week, you would place a check (✓) in book 3 under rate of occurrence.

Rater response determines how you respond to different behaviors as they occur within the classroom setting. For example, you may respond to a behavior such as not paying attention with a warning glance. On the other hand, you may respond to fighting by temporarily removing the child from the classroom setting. Under rater response, you are asked to indicate how you respond to different behaviors as they occur within the classroom by indicating which of the techniques under rater response you typically use in coping with the behaviors listed in this scale. It is recognized that you use different techniques with the same behavior, depending upon the situation; but you are asked to indicate which technique you usually or typically use in coping with the behavior in question.

Behavioral effects indicates how disruptive the behavior is to the classroom atmosphere. Some behaviors are very disruptive of a learning climate while other behaviors are minimally disruptive.

3. Rate of items in the first part of the scale as follows: If you have observed a particular behavior in the classroom, place a check (✓) in the appropriate boxes after that item. If you have not observed a given behavior in a child, place a check in the (0) box under rate of occurrence and leave the other two sections (Rater response and Behavioral Effects) blank for that item. In the second part of the scale, simply indicate the frequency with which behaviors occur that you have observed. Read all items carefully and respond to every item in the scale.

4. Indicate your judgments in each of the three scoring areas according to the following criteria.

Section A: Rate of Occurrence

(0) The behavior has never occurred.
(1) The behavior occurs at least once every two months.
(2) The behavior occurs at least once a month.
(3) The behavior occurs at least once a week.
(4) The behavior occurs at least once a day.
(5) The behavior occurs at a constant or near constant rate, (more that once a day).

Section B: Rater Response: When this particular behavior occurs, do you

(1) Ignore the behavior?
(2) Give the child a warning glance?
(3) Interact verbally or physically with the child?
(4) Temporarily remove the child from the classroom setting?
(5) Refer the child to an outside source, i.e., counselor, phychologist, or separate referral agency?

Section C: Rater Reaction

(0) Does not apply.
(1) The behavior is not disruptive.
(2) The behavior is slightly disruptive.
(3) The behavior is moderately disruptive.
(4) The behavior is disruptive.
(5) The behavior is very disruptive.

5. Enter appropriate criticisms about the design, item wording, format, and/or directions of this instrument.

_____

_____

_____

_____

_____

6. Sample item:

	Section A Rate of Occurrence	Section B Rater Response	Section C Behavioral Effect
	0 1 2 3 4 5	1 2 3 4 5	0 1 2 3 4 5
1. Shouts back when corrected in class.		✓	✓

This behavior is rated as: occurring at least once a month; the rater ignores the behavior; the behavior is moderately disruptive of a learning climate.

PART ONE

	Rate of Occurence	Rater Response	Behavioral Effects
	0 1 2 3 4 5	1 2 3 4 5	0 1 2 3 4 5
1. Does not obey commands or directives.			
2. Terminates an irritating or inappropriate behavior if verbally reprimanded, only to resume the behavior when he is not being observed.			
3. Creates a distrubance during class activities in which he is not interested or skilled.			
4. Pouts.			
5. Does not play in games with other children.			
6. Does not attend to a given task when asked to do so.			
7. Makes verbal statements such as: You can't make me do this!			0 1 2 3 4 5 . 1 2 3 4 5 0 1 2 3 4 5
8. Refuses to do any school work for a period of time.			
9. Attempts to yell the teacher down in front of the class.			
10. Argues and demands the last word.			
11. Leaves the classroom without permission.			
12. Refuses to perform or speak before the group when requested.			
13. Screams, bangs objects when denied something.			
14. eeds to do things before ructions are finished.			

	Rate of Occurrence	Rater Response	Behavioral Effects
15. Does not follow rules of games, class activities.	0 1 2 3 4 5	1 2 3 4 5	0 1 2 3 4 5
16. Does not mind or obey until physically punished.			
17. Protests about changes in his routine.			
18. Requires control from others before conforming to limits.			
19. Ignores warnings and reprimands.			
20. Encourages destructive activity or disobedience in others.			
21. Comments that he hates school.			
22. Displays violent temper tantrums.			
23. Engages in fights on the playground.			
24. Does not enter into relationships with other children.			
25. Makes lewd gestures.			
26. Shouts back when corrected in class.	0 1 2 3 4 5	1 2 3 4 5	0 1 2 3 4 5
27. Manipulates other children in order to get them to do what he wishes.			
28. Does not follow directions given by the teacher but will follow directions contained in a textbook or assignment.			
29. Tattles on other children.			
30. Makes contrary to fact statements.			
31. Threatens to kill others.			
Teases other children.			

PART TWO

		Rate of Occurrence
33.	Starts many activities, but does not finish them	0 1 2 3 4 5
34.	Uses his hands in a clumsy fashion.	
35.	Does not initiate conversations with other children.	
36.	Withdraws when teased by other children.	
37.	Apologizes for himself/his behavior.	
38.	Utters non-sensical phrases or sentences.	
39.	Expresses worry or concern about bad grades, health, etc.	
40.	Drops an activity when he loses at that activity.	
41.	Distracted from the task at hand by ordinary classroom stimuli, minor noises, movements, etc.	
42.	Loses interest in what he is doing and begins to disturb the class.	0 1 2 3 4 5
43.	Does not take his turn in group activities.	
44.	Prefers to play with younger children even though children his own age are available.	
45.	States others are to blame for his actions.	
46.	Tells stories which exaggerate the truth.	
47.	Volunteers for classroom status assignments but does not finish them.	
48.	When presented with a task, withdraws from the situation.	
49.	Writes phrases in an immature fashion using [illegible]ge and badly formed letters.	

Rate of Occurrence

50. In structured physical activities, refuses to be a team leader if chosen for the position. 0 1 2 3 4 5

51. Requests praise or approval for tasks attempted.

52. Does not ask for directions to be repeated even when it is obvious he does not understand them.

53. Mimics speech of others.

54. Talks out of turn.

55. Shifts from one activity to the next without accomplishing either.

56. Comments that he is tired.

57. Stumbles or falls

58. Must have things in perfect order. 0 1 2 3 4 5

59. Seeks approval from teacher for tasks attempted.

60. Comments that he is unable to complete a required classroom activity.

61. Answers questions about himself with "I don't know" or fails to answer.

62. Does not engage in group activities on the playground.

Scale II

BEHAVIOR RATING SCALE

Demographic Information:

Name of Pupil_____ Date of Birth_____

School_____ Grade_____

Sex of Rater_____ Sex of Pupil_____

Name of Rater_____ Date_____

Instructions to the rater:

1. This scale is designed for the purpose of identifying behaviorally disturbed children. Items in the scale represent OVERT BEHAVIORS WHICH CAN BE VERIFIED BY OBSERVATION. Thus, if you have not observed a particular behavioral item in the classroom, you would indicate in the scoring section that the behavior had never occurred.

2. In the first part of the scale, three rating judgments are required for each behavioral item: (a) rate of occurrence, (b) rater response, (c) behavioral effects. One judgment is required under (a) rate of occurrence; one judgment is required under (b) rater response; and one judgment is required under (c) behavioral effects. Thus, there would not be more than three rating judgments per item.

   Rate of occurrence is designed to secure information on the frequency with which a particular behavior occurs within the classroom setting. For example, if a behavior occurs one or more times in a week, you would place a check ( ) in box 3 under rate of occurrence.

   Rater Response determines how you respond to different behaviors as they occur within the classroom setting. For example, you may respond to a behavior such as not paying attention with a warning glance. On the other hand, you may respond to fighting by temporarily removing the child from the classroom setting. Under rater response, you are asked to indicate how you respond to different behaviors as they occur within the classroom by indicating which of the techniques under rater response you typically use in coping with the behaviors listed in this scale. It is recognized that you use different techniques with the same behavior, depending upon the situation; but you are asked to indicate which technique you usually or typically use in coping with the behavior in question.

   Behavioral effects indicates how disruptive the behavior is to the classroom atmosphere. Some behaviors are very disruptive of a learning climate while other behaviors are minimally disruptive.

3. Rate of items in the first part of the scale as follows: If you have observed a particular behavior in the classroom, place a check( ) in the appropriate boxes after that item. If you have not observed a given behavior in a child, place a check in the (0) box under rate of occurrence and leave the other two sections (Rater Response and Behavioral Effects) blank for that item. In the second part of the scale, simply indicate the frequency with which behaviors occur that you have observed. Read all items carefully and respond to every item in the scale.

4. Indicate your judgments in each of the three scoring areas according to the following criteria.

Section A: Rate of Occurrence

(0) The behavior has never occurred.
(1) The behavior occurs at least once every two months.
(2) The behavior occurs at least once a month.
(3) The behavior occurs at least once a week.
(4) The behavior occurs at least once a day.
(5) The behavior occurs at a constant or near constant rate, (more than once a day).

Section B: Rater Response: When this particular behavior occurs, do you

(1) Ignore the behavior?
(2) Give the child a warning glance?
(3) Interact verbally or physically with the child?
(4) Temporarily remove the child from the classroom setting?
(5) Refer the child to an outside source, i.e., counselor, psychologist, or separate referral agency?

Section C: Rater Reaction

(0) Does not apply
(1) The behavior is not disruptive
(2) The behavior is slightly disruptive.
(3) The behavior is moderately disruptive.
(4) The behavior is disruptive.
(5) The behavior is very disruptive.

5. Enter appropriate criticisms about the design, item wording, format, and/or directions of this instrument.

_____

_____

_____

6. Sample item:

	Section A Rate of Occurrence						Section B Rater Response						Section C Behavioral Effect					
	0	1	2	3	4	5	1	2	3	4	5	0	1	2	3	4	5	
1. Shouts back when corrected in class.																		

This behavior is rated as: occurring at least once a month; the rater ignores the behavior; the behavior is moderately disruptive of a learning climate.

	Rate of Occurrence	Rater Response	Behavioral Effects
	0 1 2 3 4 5	1 2 3 4 5	0 1 2 3 4 5
1. Willingly accepts challenges and gets into fights.			
2. Goes through other children's possessions without authorization.			
3. Responds to teasing with physical aggression			
4. Provokes other children in the classroom by disturbing, teasing or shoving them.			
5. When angry, slams books on the desk, slams doors, kicks chairs, etc.			
6. Uses profane language in the classroom.			
7. Initiates fights with other children.			
8. Comments that he hates his teacher.	0 1 2 3 4 5	1 2 3 4 5	0 1 2 3 4 5
9. If the teacher insists that he do school work when he has refused, throws a temper tantrum, cries, screams, etc.			
10. Provokes fights on the playground, reports on fight, then denies having initiated the fight.			
11. Will destroy or take apart something he has made rather than show it or ask to have it displayed.			
12. Threatens other children with physical violence.			
13. Attacks other children with potentially dangerous objects: knives, pencils, sharp objects, etc.			

	Rate of Occurrence	Rater Response	Behavioral Effects
	0 1 2 3 4 5	1 2 3 4 5	0 1 2 3 4 5
14. When angry, will destroy his own possessions: books, models, pencils, paper, etc.			
15. Refuses to recognize the fact when he is proven mistaken or wrong.			
16. Threatens to call in his parents to extricate himself from a hostile interaction with the teacher.			
17. Makes loud verbal outburst without raising his hand and securing permission to speak.			
18. Cries when things do not go his way.			
19. Steals things from other children.			
20. Destroys or defaces property other than his own.			
21. Forces the teacher to give him her attention.			
22. Refuses to recite aloud in class.	0 1 2 3 4 5	1 2 3 4 5	0 1 2 3 4 5
23. Does not express himself orally			
24. Strikes another child and then leaves, not staying to carry on with the other child.			
25. Interrupts other children while they are working.			
26. Pesters other children.			
27. Imitates the behavior of his classmates in a mechanical fashion.			
28. Asks to be excused from activities in which he is required to participate.			

	Rate of Occurrence	Rater Response	Behavioral Effects
	0 1 2 3 4 5	1 2 3 4 5	0 1 2 3 4 5
29. When mistreated by other children, takes out his frustrations on another inappropriate person or thing.			
30. Corrects other children.			
31. Picks on smaller or weaker children.			
32. Tries to settle disagreements aggressively, e.g., by bullying or yelling.			

PART TWO

	Rate of Occurrence
	0 1 2 3 4 5
33. Complains of headaches, cramps general body aches.	
34. Does not respond to verbal inquiries or questions from the teacher.	
35. Hesitates a long time before making choices.	
36. If not working well at the task assigned, drifts off and finds a way to comfort himself.	
37. Stutters.	
38. Comments that nobody likes him.	
39. Is absent from school when a major assignment or test is due.	
40. Appears tired and lethargic even though not suffering fatigue from physical activity.	
41. Remains in one position for long periods and stares fixedly while doing so.	
42. Shows muscle irregularities, spasticity, rigidities.	
43. Comments that a particular activity is too hard for him and then quits.	

Rate of Occurrence

44. Comments that a particular activity is too hard for him and then quits. 0 1 2 3 4 5

45. Does not pronounce words clearly.

46. Interrupts the class with comments which have no bearing on the class activity.

47. Repeats same acts over and over in a mechanical fashion.

48. Comments that he is stupid.

49. Complains of difficulty in breathing.

50. Cries without apparent provocation.

51. Comments that he does not feel well.

52. Is easily thrown off and makes errors.

53. Complains of others' unfairness toward him.

54. Although he does not create a disturbance or disrupt the class, does not do any school work for given periods of time.

55. Is hyperactive; e.g., constantly moving.

56. Gives excuses for not getting work in on time.

57. Cries whenever the teacher directs attention toward him.

58. Reports difficulty in thinking; e.g., I can't concentrate.

59. Uses baby talk.

60. Talks to himself.

61. Comments that others are out to get him or have it in for him.

62. Displays poor coordination in physical activities.

Stage Three

Behavior Observation Form

Re-programing Project
Oregon Research Inst.
September, 1968

OBSERVATION IN THE SCHOOL: DESCRIPTION OF A CODING FORM

R.S. Ray, D.A. Shaw, G. R. Patterson

The following is a trial-and-error refinement of a school observation technique used by the project's observers during the past year. It is a method of "characterizing" school situations for a given child in such a way as to facilitate understanding the determinants and consequences of social behaviors as well as the relationship of those behaviors to the classroom setting.

Each coding sheet represents six minutes of behavior for a given subject. The "deviant" child may be compared to his "normal" peer by alternating the two as subjects of observation. (We conventionally observe the "deviant" child for 12 minutes (2 coding sheets) and then select a "normal" peer at random to observe for 6 minutes (1 coding sheet) before returning to the "deviant" child). Each coding sheet provides the following information: behavior of the subject, social consequence, agent supplying consequence, and description of the classroom situation.

The rating form is set up as a grid. Each horizontal line in the grid represents a fifteen-second time interval. The grid is divided into two-minute "chunks" simply for the convenience of the observer in reading the behavior codes. Using the observation clipboard set for fifteen-second intervals, the observer moves down one line each time he receives a signal from the clipboard, i.e., at the end of each fifteen-second time block. (If no clipboard-timer is available, a stopwatch or school clock will generally suffice.) The vertical spaces in the grid correspond to the behaviors listed at the top of the two-minute section. During each fifteen

second interval the observer records both the behavior of the subject and the social consequences of his behavior by placing the appropriate response and agent codes in the space beneath the appropriate behavior code. The response codes and the agent codes are listed at the top of the coding sheet. For example, if the subject is not attending to the teacher's explanation of the lesson and the teacher "calls him down" (i.e., disapproves) the interaction would be coded as follows:

Depending upon the rate at which things are happening in the classroom, the observer may code more than one subject behavior and more than one consequence during each fifteen-second interval. In most cases, however, there will be only one primary behavior or interaction. It is not necessary to make more than one coding entry for behavior which continues unchanged throughout the fifteen-second interval. The subject's behavior during the fifteen-second period should be "characterized" by the coding; it need not be described in sequential detail. The observer should check the situation category at the right side of the coding sheet which best describes the situation during each 2-minute section. The categories are as follows:

Classroom:

Group: To be used whenever the classroom activity is essentially group rather than individual work, e.g., teacher presenting lesson to entire class, subject in a reading group, etc.

Individual: To be checked whenever the subject is involved in individual rather than group work, e.g., sitting at desk doing arithmetic, reading; not listening as a group to teacher or working together in a group.

Transition: This category will generally be checked when the entire group is switching from one kind of activity to another, e.g., lining up to go to assembly, moving chairs to form a reading group. Frequently is movement associated but it would not be considered inappropriate in this situation, e.g., entire class moving from desks to reading groups would be coded MO (appropriate group behavior) rather than MO (movement).

Recess: Checked during any regular recess period whether child is outside or remains in classroom. A short description of the child's activity should be written on the lines provided at the right of each two-minute section (this applies to classroom as well as recess behaviors). Generally during recess periods, the observer will be primarily interested in coding either PL (playing with others) or AL(isolated from others) and possibly AG (aggression); other behavior codes are not so relevant to the usual recess activities.

Description of Codes:

CLASSROOM BEHAVIORS:

NY (noisy): Coded whenever the subject is talking loudly, yelling, or making other deliberate, inappropriate noise (such as banging books or scraping chair back and forth) which is actually or potentially disruptive to others.

AG (aggression): Actual or attempted physical abuse of another, e.g., John hits Bill or John starts to hit Bill and is stopped by the teacher. This includes pushing, shoving, threatening, bossy.

NA (not attending): Subject is not attending to his work or to a lesson being taught, etc.; may be looking out the window, watching the observer or other children, drawing when he is supposed to be watching teacher demonstrate arithmetic, leaning down to tie his shoe, turning in his chair.

PI *(peer initiation): Peer talks to, pokes or in some way trys for attention of S.

IP *(Initiation to peer): S talks to, or in some ways trys for attention of peer.

MO (movement around room): Coded whenever subject is moving around room (other than times when entire group is moving, as in transition periods); observer need not try to decide whether each movement is appropriate or inappropriate; that question is better decided by comparing rate of movement across subjects.

IW *(Inappropriate task): Work on task not assigned or specified by teacher for that time.

NO (appropriate group behavior or normative behavior): Coded whenever the subject's behavior is task-directed activity which is appropriate for that time and situation. Included would be listening to the teacher explain a lesson, painting during an art class, singing with others during music, lining up with the rest of the class to go out for recess, etc. The observer should take care not to include any behavior which might be more appropriately characterized as recitation.

WK (work): A child may be engaged in appropriate group activity but not working e.g., observe a movie. Work means at desk on academic projects. Must work on teacher assigned task. Record when engaged in reading, writing, arithmetic, basic skills.

RE (recites): Coded whenever subject recites, answers a teacher's questions, reads out loud, gives a speech, or performs before the class.

VO (volunteers): Coded whenever subject raises his hand or in some other manner indicates a desire to recite or do whatever else the teacher may have asked for, e.g., someone to pick up papers; may be either in a class discussion or in a small group.

TI (teacher initiation): Coded when the teacher "calls on" the subject or comes to his desk or activitiy area to speak to him; this interaction must be initiated by the teacher and not be a response to an initiation by the subject.

IT (initiation to teacher): Coded when the subject indicates that he wants some attention from the teacher; he may raise his hand, speak or go to her; this behavior is differentiated from "volunteer" in that the subject's initiation to the teacher is not in conjunction with class discussion, group study, or reciting.

RECESS BEHAVIORS ONLY:

PL (play with others): Coded when the subject is clearly in the company of others, whether talking, playing a game or just walking around.

AL (alone, isolated from others): Coded when the subject is engaged in solitary activity, whether playing a game or not.

*These three categories were added or expanded by the present authors (Walker and Buckley) for use in ELP.

RESPONSES:

O (no response): Coded when no response from teacher, peers, or observer follows a behavior. There may be no response because the behavior does not demand one or because the subject is clearly being ignored, the observer should note this at the side rather than attempt to code the response any differently).

A (attention): Coded whenever the agent listens to or looks at the behaving subject; this is a neutral kind of response with no obvious approval or disapproval in the attending response.

P (praise): Coded when the subject receives praise or approval from an agent; may be verbal behavior or consist of gestures, e.g., smiles, head nods, applause.

C (compliance): Coded when subject complies with a command from another.

NC (non-compliance): Coded when subject does not comply with a command from another. Neither C or NC will probably be used very often; if they are, they will probably be responses of the subject to TI.

PH + or - (physical contact positive or negative): Positive physical contact would include such behaviors as hugs, pats on the back; negative physical contact would include aggressive behaviors from an agent such as hitting, spanking, etc.

D (disapproval): Coded when a subject behavior is followed by verbal or gestural disapproval from an agent; examples might be frowning, negative head nods, "you shouldn't have done that," etc.

AGENTS:

Each response should be subscripted with one of the following:

T (teacher), P (peer), O (observer)

to indicate the agent of the response. O should rarely have to be used.