DEVELOPMENT OF A BEHAVIORAL RATING SCALE FOR
PROFOUNDLY RETARDED CHILDREN

APPROVED:

Mark E. Bonney
Major Professor

Earl W. Hoover
Minor Professor

Dwane Hingery
Director of the Department of
Education and Psychology

Robert B. Toubour
Dean of the Graduate School
DEVELOPMENT OF A BEHAVIORAL RATING SCALE FOR
PROFOUNDLY RETARDED CHILDREN

THESIS

Presented to the Graduate Council of the
North Texas State University in Partial
Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

BY

Paul A. Rhoads, B.A.
Denton, Texas
August, 1968
# TABLE OF CONTENTS

LIST OF TABLES ........................................ iv

Chapter

I. INTRODUCTION ........................................ 1
   The Purpose of the Study
   Problem of Study
   Related Literature

II. METHODS .............................................. 8
   Subjects
   Procedure
   Treatment of Data

III. RESULTS ............................................. 14
   Items Chosen for Scale as a
   Result of Observations
   Statistical Results

IV. SUMMARY OF FINDING ............................... 23
LIST OF TABLES

Table                      Page

I. Reliability Coefficients Between Trainers' Ratings       .20
II. Average Scores for Major Groups and for Levels Within Groups .21
CHAPTER I

INTRODUCTION

The Purpose of the Study

There are an estimated five million mentally retarded children and adults in the United States; of these sixty thousand are classified as severely and profoundly retarded. (7, p. 525). This group never attains a mental age greater than that of a three year old child. Differentiation and rating of the profoundly retarded are extremely difficult since IQ testing of this group is drastically limited and in most instances impossible. In a review of literature related to intelligence testing of the profoundly retarded, not one article could be found relating either to the testing of the profoundly retarded or to the devising of an IQ scale.

A scale for rating a group of profoundly retarded children is needed by institutions for retarded children, for several reasons. In the first place, it is needed as a means of discovering if the child is profoundly retarded or if he should be placed in a higher program from which he would receive more benefit. In the second place, such a scale would help to determine a child's degree of profound retardation. This would

*The American Association for Mental Deficiency's classification was used in this study. (Heber, 1959)
help in determining where a child should be placed within the group classified as "profound" and would help in determining the nature of programs to which the child is exposed. Finally, a developmental rating scale for the profoundly retarded would be useful as a means of determining a child's progress over a period of time.

Since the psychometric testing is not applicable to this level of retardation, differentiation of the profoundly retarded will have to be done on the basis of individual behavior rather than test results. Kazue, in pointing out how overt behavior is being used more and more to measure human potentials, states that "behavior is progressing toward acceptance as a dimension with measurement potential" (11, pp. 1-6). However, only two behavioral scales exist at the present, one of which is the Vineland Social Maturity Scale. Although the Vineland is the oldest, it has serious limitations. It was published by Edward A. Doll in 1953 and was originally intended only as an experimental method (8, 9, 10). It was loosely standardized on normal children and adults living in Vineland, New Jersey, not retardates. These data obtained in one town during the mid-thirties probably have limited applicability to the general population today. No further normative data on this scale have been published despite Doll's expectations that the scale would be revised and restandardized (15, p. 442). In a review of the literature only a few articles could be found relating to the use of the Vineland for
determining the level of mental retardation, and none of these related to the testing of the profound level. The practice of classifying behavior of the mentally retarded according to the incomplete normative data of the Vineland scale is very dubious (15, p. 444).

The other behavioral scale, the Cain-Levine Social Competency Scale, is a recent one (2). It was standardized on trainable mentally retarded children, all of whom were living at home; therefore, it is lacking in adequate standardization.

The need for a behavioral scale that can differentiate between the lower levels of mental retardation is being expressed by a growing number of writers (3, 4, 5, 6, 11, 12, 13, 14, 16, 17, 18). Bathazus has pointed out that "with the more severely retarded resident, representative behavior appears to be more significant from an evaluative and treatment point of view than measurement of deviance from the norm" (1, pp. 16-19).

Problem of the Study

The purpose of this study was first to determine what criteria could be used on a behavioral scale to evaluate the profoundly retarded, and then to acquire some data bearing on the reliability and validity of a scale based on such criteria.
Related Literature

There are few related studies in this field. The first such study (3, 4, 5, 6) was conducted by Clark in 1932. However, Clark derived his study from a strictly psychoanalytic point of view. He organized his observations in terms of Freudian concepts. His study has also been criticized for lacking adequate evidence to back up his results.

Woodward (16; 17, p. 98) applied Piaget's theory of intellectual development to the study of the mentally retarded and through observations of their performance on six tasks (not set up by the experimenter) concluded that their behavior was characteristic of much younger children. It should be added, though, that a profound level of retardation was not included in this study.

Wright attempted to differentiate between retardates, schizophrenics, and normal children, using a Q-sort of five dimensions of behavior. He was successful in differentiating the schizophrenic group but not the retarded group. He concluded that "a second index is needed which will separate the retarded from the normal" (18, pp. 169-185).

A final study, which closely parallels the present study, was made by McKinney (13), in which a group of severely retarded children was observed twice a day for periods of five minutes, for twelve days. During these observations the children were rated on a behavioral list of 109 items. Eighty items were extracted as being indicative of the severe
level of retardation. However, a profound group of retardates was not used in this study, nor were any further validation studies conducted on these items.
CHAPTER BIBLIOGRAPHY


8. Training School, The Vineland Social Maturity Scale, Vineland, New Jersey, 1936.


CHAPTER II

METHODS

Subjects

The subject for this study were chosen at random from the general population of students on one side of two dormitories of mentally retarded children at Denton State School, Denton, Texas. One dormitory was classified as "severe", with IQ's ranging from twenty to thirty-five, and the other as "profound", with IQ's estimated below twenty.

In the initial behavioral observations, thirty subjects were used, the students had previously been divided into three levels of impairment in each dormitory by the school psychologist. Placement of children on these levels was made by the school psychologist, based on his observations and reports from co-workers. Five subjects were drawn from each level in both dormitories, using a total of thirty subjects in all. In the second half of the study a total of seventy-four subjects were used from the two dormitories.

Procedure

The subjects were observed in their dormitories while involved in their normal daily activities. The observer had minimum interaction with the subjects. Observations were made
on only one subject at a time. All actions by the subject and all actions by others which directly affected the subject were written down by the observer.

A time sampling technique was used in which each subject was observed for a five minute period. Each one was observed a total of twenty times over a three-week period. A schedule was devised for observing the subject to insure a wide sampling of his behavior. The observations began with the first subject selected from the highest group of the "severe" dormitory. At the end of five minutes a second subject was observed. This continued until all five subjects in that group had been observed. The observer then moved again to the first child chosen from that group and continued observation until all five had been observed. The same procedure was used with the lowest group in the "severe" dormitory. When observations were completed on this group, the observer proceeded to the "profound" dormitory and observed the chosen subjects in the same manner.

The times of observations were broken into eight segments of two hours, each beginning at 6 a.m. and ending at 10 p.m. A check was kept to insure that each group was observed two times in each segment. The extra times were used to catch the subject in some particular activity which had been missed. The observations began when the subjects awakened and ended when they were asleep.
As a result of these observations significant behavior was selected from the records of the subjects which differentiated between the various levels. Those aspects of behavior which seemed to divide the various levels were chosen. Two criteria were used: first, those items of behavior which a higher group could do and a lower group could not were chosen; second, those items of behavior in which higher groups had greater proficiency were selected.

Altogether, 600 five-minute units of behavior were recorded in the course of the observations. These time-units included approximately 6,300 individual behavior units. Each individual description of the behavior of the subject constituted an individual unit. Several examples of notes taken in one five-minute segment are given below:

1. Subject played with toy
   dropped toy and picked up another
   banged toy against table
   picked another toy up with left hand
   hit two toys together
   shaking two toys together
   dropped toy from left hand
   picked up new one
   stood up with toys in each hand
   told by trainer to sit down

2. Made vocal grunts while watching trainer
   watched trainer play with another boy
   rocked chair back and forth for about a minute
   stood up
   watched trainer leave room
   made nasal whining noise
   tried to follow trainer
   was turned back by trainer and told to go back to seat
   sat down in floor
   started crying
   began hitting head with hand
3. Sat on floor by trainer
rocked head from side to side
lay over on side
fell asleep
was stumbled over by another boy ---
did not awaken

The extracted results were then organized into a scale by the experimenter. Each item on the scale represented those behavioral units which differentiated the two groups. Items which the higher group did over 50 per cent of the time and the lower did less than 50 per cent of the time were extracted. These items were organized into nine general headings: performance, eating, toilet training, vocabulary, memory, dressing, attention, and response to frustration.*

The scale was then distributed to twelve trainers who rated the children in their groups according to items on the devised scale. The trainers were instructed to give the subject credit if the item applied to him a majority of the time. Each item was worth one point; a total of forty-five points was possible. The trainers worked with the children five to six days a week for eight hours a day. Six groups of subjects were used: three profound groups and three severe groups. Each group was rated by two trainers who worked with the children. In all, seventy-four subjects were rated, thirty-nine severe and thirty-five profound.

*Those headings and the individual units which make them up will be discussed at length in Chapter III.
Treatment of Data

The results were then statistically treated in two ways. First, an inner-rater reliability coefficient was obtained for the amount of agreement between six pairs of trainers' ratings of the six groups. Inner-rater reliability coefficients were also obtained for each of the six pairs of trainers rating each of the six groups.

Then the average between the two groups, "severe" and "profound", was subjected to a t test to determine the amount the scale was able to differentiate between the two levels. One of the two trainers who rated each group worked with the group six days; the other, three days. The ratings given by the trainer who worked with the group six days were used in the t test.

An analysis of variance was then worked out for the various levels of the two groups to determine if the scale differentiated at the finer levels: between the upper and middle and the middle and the lower of the severe group, between the lower of the severe group and the upper of the profound group, and between the upper and the middle and the middle and the lower of the profound group.
CHAPTER III

RESULTS

Items Chosen for Scale as a Result of Observations

A total of nine areas were found by which differentiation could be made between the profound and severe groups: performance, eating, toilet training, communication, vocabulary, memory, dressing, attention, and response to frustration. Each of these areas was broken into behavior specifics on the scale by the experimenter.

Response to frustration was discovered as a differentiating factor in this study. It had not been mentioned in earlier studies. Children at the profound level of mental retardation will perform acts of self aggression for no apparent reason or when strongly frustrated, whereas a child severely retarded rarely, if ever, responds to frustration in this manner. The behavior specifics as they appeared on the scale are as follows:

RESPONSE TO FRUSTRATION

Does not perform acts of self aggression for no apparent reason, that is without frustration. (does not perform acts of self aggression such as biting hand, hitting self, hitting head against something for no apparent reason).

Even when frustrated seldom responds with self aggression, (when frustrated by another person seldom if at all responds with self aggression such as biting hand, hitting self, hitting head against wall).
When frustrated does not respond at all with self aggression (when frustrated by being restrained, forced to do something, hit by another subject, responds in some other way than with self aggression).

A second area, not previously noted, is that of attention. It was found that in the most profoundly retarded group the subject's attention could be gained only by physical contact, whereas at the highest level the subject's attention could be drawn and held by an event which was not in his environment. It was also noted that the most profoundly retarded would make random gestures, body movements or waving of the head, with no apparent purpose. The following six behavior specifics related to attention were included:

**ATTENTION**

Attention can be drawn and held by event which physically affects subject (if hand is taken, or is pushed or hit by another subject, will turn attention toward the stimulus).

Attention can be drawn and held by events in his immediate environment which directly affect him (action directly involves the subject, such as someone talking directly to him, can draw and hold his attention).

Attention can be drawn and held by events in immediate environment which are not of a physical nature and do not directly affect him (if another child close by cries, trainer or parent talks to someone close by, or something is knocked over, subject will fix attention on it).

Attention can be drawn and held by event which is not in subject's immediate environment (someone entering the room, the TV, a child at some distance cries, or some event takes place which is not in the subject's immediate environment, subject will fix attention on it).

Does not make meaningless random gestures or movements (does not make meaningless, random, repetitive waving, rocking or turning movements with hands, feet or body)

Six items were included in the area of dressing:
DRESSING

When dressed will keep clothes on (will not take clothes off after dressing or being dressed).

Can put clothes on though may require assistance (puts clothes on when told but may require assistance does not have to button, zip zippers, or manipulate snaps, gets credit even if clothes are put on incorrectly).

Can dress without assistance (puts clothes on without assistance, may put on incorrectly and still get credit, but does not have to button, zip zipper, or manipulate snaps).

Can dress self correctly without assistance (puts clothes on correctly without assistance but does not have to be able to manipulate snaps, button, or zip zipper).

Can manipulate zipper and/or snaps (can work zipper and snap clothes).

Can button clothes (Can button own clothes though may require considerable time to do so)

Memory included items relating to the student's ability to remember from past experience such things as finding his way around the dormitory and responding to his own name. The following behavior specifics were included in the scale.

MEMORY

Responds to familiar face (when familiar dormitory parent, trainer, teacher or parent approaches gives some sign of recognition).

Responds to name when called (will turn head or move in appropriate direction when name is called).

Knows way around dormitory or home (can find bathroom, bed, or any familiar place)

Can remember way around dormitory or a new seating arrangement with only one or two exposures (with only one or two experiences in a new setting can remember correct places to go).
Vocabulary was not necessarily speaking vocabulary but evidence of an association between a word and an object. At the lower levels several of the subjects did not have a speaking vocabulary but they could point to an object or part of the body if named by someone else. Subjects at the higher levels, however, did possess a speaking vocabulary. The subject's ability to follow verbal commands was also included in this section. The behavior specifics in this area were these:

**VOCABULARY**

Knows names of few objects or persons in immediate environment (student must indicate that he forms an association between stimulus word and object, credit is given either if student names object or if someone else names object and student can point to it).

Follow simple commands (makes appropriate responses to such commands as "sit up", "come", "sit down").

Knows names of large number of objects or persons in environment (subject must indicate that he forms an association between stimulus words and object, credit is given either if subject names object or if someone else names object and subject can point to it, however, there must be an extensive number of objects).

Can use simple words appropriately, with understanding (has a speaking vocabulary of a few words and can use them appropriately, words do not have to be used in combination appropriately).

Can follow complicated commands (makes appropriate responses to such commands as "go to bed", "go get the chair", "give me the ball").

Can use simple word combinations appropriately, with understanding (can use two or more words in combinations appropriately).

Can speak appropriate simple sentences (can use several words in combination to make appropriate simple sentences).
Communication could also be termed interaction. Included here was the extent to which the subject attempted to communicate with someone. The behavior specifics here were as follows:

**COMMUNICATION**

Direct vocalizations and/or gestures at someone (vocalizations and gestures may be with meaning or without but are directed at a specific person)

Attempts meaningful communication, verbal or non-verbal with authority figure (communication could be verbal or in form of gestures but must be directed at some authority figure: dorm parent, trainer, parent, and must be to some extent meaningful, such as pointing to pants when wants to go to bathroom).

Attempts meaningful communication with peer group, either verbal or non-verbal (communication can be either verbal or non-verbal but must be meaningful and whether verbal or in form of gestures directed at fellow students).

Students at lower levels of retardation did not attempt to control their bowels or bladder and would soil in their clothes. However, at higher levels the students would go to the bathroom on their own. The following three items were placed in this category on the scale:

**TOILET TRAINING**

When on toilet performs necessary functions (credit is given if when on toilet performs necessary functions, even if necessary to place there or has tendency to soil if not placed there in time).

Will go to toilet on own initiative (will go to the toilet when it is necessary but may occasionally soil).

Will go to the toilet on own initiative and rarely if ever soils (rarely if ever soils).
The largest area, that of eating, including both eating and drinking. The three items on the scale in the area of drinking and seven in the area of eating were these:

**EATING**

Drinks from glass (gets credit even if glass has to be held for the subject).

Drinks from glass unassisted (holds glass and drinks without aid).

Drinks from glass and only rarely spills any on self (holds own glass and does not spill any either in process of drinking or in process of getting glass to mouth).

Does not eat inedible substances (is able to discern between food and non-food substances).

Chews food before swallowing (does not swallow foods whole).

Feeds self (does not have to be fed, credit is given even if hand is used and not spoon).

Eats with spoon at least part of time (eats with spoon, credit is given even if fingers are used some of the time).

Does not try to put too much in mouth (takes appropriate bites and spoonfuls, does not put too much in mouth at one time).

Eats exclusively with spoon (only rarely, if at all, uses hands).

Eats with fork (may use spoon and fork in combination and get credit).

Performances included four behavior specifics which could actually be broken into two areas, those tasks with supervision and those without supervision.
PERFORMANCE

Can perform simple tasks (pick up paper, toys, carry food tray, moving chairs but may require supervision in doing so).

Can perform more involved tasks (mopping floor, sweeping, moving light furniture, credit is given even if performance is crude or supervision is required).

Can perform simple tasks without supervision (pick up paper, toys, carry food tray, move chair, may have to be told to do so but does not require supervision).

Can perform more involved tasks without supervision (mopping floor, sweeping, cleaning up, may have to be told to do so but does not require supervision).

STATISTICAL RESULTS

Using these items the trainers rated each subject by giving him a check mark if the item applied to him and leaving it blank if it did not. Each item was worth one point. The points were added up to equal a total score for each subject. A maximum score of forty-five was possible.

With these items forming the scale, a reliability study was conducted to determine the amount of agreement that was obtained between two raters rating the same group of children (see Table I).

Each of the six groups of retarded children, three profound and three severe, used in the reliability studies of the devised scale, was rated by two trainers, yielding six reliability coefficients for the amount of agreement between the trainers' ratings. The obtained coefficients ranged from .77
to .93*, as shown in Table I, with an average of .87. This would indicate that the behavioral scale possesses a high degree of inter-rater reliability.

**TABLE I**

**RELIABILITY COEFFICIENTS BETWEEN TRAINERS**

<table>
<thead>
<tr>
<th>Group</th>
<th>Level Within Group</th>
<th>Number of Subject</th>
<th>Inter-rater Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Highest</td>
<td>13</td>
<td>.97</td>
</tr>
<tr>
<td>S</td>
<td>Middle</td>
<td>12</td>
<td>.98</td>
</tr>
<tr>
<td>S</td>
<td>Lowest</td>
<td>14</td>
<td>.83</td>
</tr>
<tr>
<td>P</td>
<td>Highest</td>
<td>11</td>
<td>.92</td>
</tr>
<tr>
<td>P</td>
<td>Middle</td>
<td>12</td>
<td>.91</td>
</tr>
<tr>
<td>P</td>
<td>Lowest</td>
<td>12</td>
<td>.77</td>
</tr>
</tbody>
</table>

It was found, as a result of a t score of 2.43 that the average scores of the two groups were significantly different at the .05 level (see Table II). This would indicate that the scale differentiates adequately between these two levels of retardation.

*The coefficient for the highest group of severes is not included in this average due to the fact that the scale did not go high enough for this group, causing both raters to give nearly maximum credit to all of the subjects in that group.*
An analysis of variance was then carried out between the various levels of the two groups to determine if the scale differentiated at the finer levels; between the upper and middle and the middle and the lower on the severe group, and the lower of the severe group and the upper of the profound group, and between the upper and the middle and the lower of the profound group. As a result, an $F$ of 1.84 was obtained, which was non-significant. The scale was, however, devised to differentiate between the profound and severe levels of retardation, which it does adequately.

**TABLE II**

**AVERAGE SCORES FOR MAJOR GROUPS AND FOR LEVELS WITHIN GROUPS**

<table>
<thead>
<tr>
<th>Major Groups</th>
<th>Average For Major Groups</th>
<th>Levels Within Major Groups</th>
<th>Average For Levels Within Major Groups</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe</td>
<td>Highest</td>
<td></td>
<td>44</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Middle</td>
<td></td>
<td>40</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>Lowest</td>
<td></td>
<td>37</td>
<td>2.6</td>
</tr>
<tr>
<td>Profound</td>
<td>Highest</td>
<td></td>
<td>27</td>
<td>4.4</td>
</tr>
<tr>
<td></td>
<td>Middle</td>
<td></td>
<td>19</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>Lowest</td>
<td></td>
<td>14</td>
<td>2.6</td>
</tr>
</tbody>
</table>
The problem involved in this study was to find behaviors of "profoundly" and "severely" retarded children which could be used to differentiate between the various levels of these two groups. These types of behavior were obtained by observations of "profoundly" and "severely" retarded children during their daily activity. These observations took place over a two-week period, during which time the subjects were observed from 6:00 a.m., when they rose, till 8:30 p.m., when they went to bed. Each subject was observed twenty times.

As a result of these observations forty-five items were found which differentiated between the two groups. Items which the upper levels did over 50 per cent of the time and the lower levels did less than 50 per cent of the time were chosen. The items were then placed in the form of a scale used by twelve trainers to rate six groups of subjects, two rating each group. The trainers were to check those items which the subject they were rating performed over 50 per cent of the time. These checks were then added up to yield a total score for each subject. An average reliability coefficient of .87 was obtained for the twelve ratings of the six groups.
A t test was also carried out to determine if the scale differentiated between the two major groups, "severe" and "profound". The test yielded a t of 2.43, which was significant at the .05 level. An analysis of variance was also conducted, to determine if the scale differentiated between the finer levels within the two major groups. This analysis of variance yielded an F of 1.84 which was non-significant. The scale does, however, differentiate between the two major groups.
BIBLIOGRAPHY

Books


Article


Doll, E.A., "Social Age as a Basis for Classification and Training," American Journal of Mental Deficiencies, 47 (April, 1942), 49-52.


Publications of Learned Organizations

