# EDUCATING THE SEVERELY MULTIPLY HANDICAPPED: MEASUREMENT PROCEDURES IN THE CLASSROOM

**Marty Mattison II and Rebecca Rosenberg**

*Direct formative measurement procedures have been developed to help teachers systematically specialize and individualize their instruction and data keeping for severely multiply handicapped students.*

*This article examines the reasons for measurement, types of measurement procedures and how these techniques may be applied by a teacher in the classroom. Advantages and disadvantages of various measurement techniques including percent, count, frequency, latency, and duration are discussed and an applied example of each tactic is given.*

As severely multiply handicapped students enter public school systems, teachers of these individuals have discovered that new technologies are necessary to map the educational progress of these students.

The severely multiply handicapped do not learn incidentally, nor rapidly, nor easily. Traditional educational methods such as inquiry, discovery and problem solving and traditional data systems (e.g. standardized tests, daily worksheets and pre- and post- test measures) have not proven successful with this population (Bricker, Bricker, Iacino, and Dennison, 1976).

Those who have succeeded in improving the functioning of severely multiply handicapped individuals have done so through procedures based on the principles of precise behavioral technology which provides the teacher with concrete information of student functioning, performance and growth (Cooper, 1974; Hall, 1971; Kazdin, 1975).

Teachers utilizing such techniques in the classroom have found that these severely multiply handicapped individuals can learn behaviors such as using speech or other communication modes, walking; feeding themselves with utensils, dressing and going to the bathroom on command, pre-academic and academic skills, and prevocational and vocational skills (Barton, 1973; Giles and Wolf, 1966; Gold, 1972; Guess, Sailor, Keogh and Baer, 1976; Karen, Eisner, Endres, 1975; O'Brien, Azrin and Bugle, 1972).

### Why Measurement is Important

Horner, Holvoet and Rinner (1976) suggest four primary reasons for systematically measuring behavior:

1) determining present level of performance prior to intervention
2) determining level of performance during intervention (effectiveness of intervention)
3) determining level of performance under differing stimulus conditions and serve as a prerequisite to establishing functional relationships between stimulus conditions and behavior
4) communicating the results of intervention to others.

Measurement of behavior helps a teacher become and remain sensitive to one's own behavior as well as the student's and other concerned persons. Effective teaching cannot occur without measurement (Horner, et al, 1976). Daily charting helps teachers see changes as they occur, and make appropriate decisions as needed rather than "I think" or "I feel" statements.

Proper measurement supports accountability. Charted data showing rate of progress will reveal whether the instructional plan is working with the student and will, in turn, suggest modifications in the plan if progress is minimal. As Lindsley (1968) once put it, "we have to care enough to chart" (p. 45).

### Types of Measurement Techniques

Several forms of measurement are utilized in the classroom with severely multiply handicapped including: counts, checklists, percent of correct and/or error, frequency, trials to criterion, latency and duration.

*Binary Count.* The easiest measurement unit employed is "binary data" (Gentry and Haring, 1976). Assessment batteries, such as the Uniform Performance Assessment System (1970) requiring simple "yes" or "no" marks, typify this particular measurement unit. Checklist and binary data are useful when measuring the occurrence or non-occurrence of a behavior. For example, a child is either wet or dry when arriving at school, does or doesn't take his shirt off independently, does or doesn't scoop from a dish. As Gentry and Haring (1976) cite: "Binary data can be used for a variety of behaviors where the relevant characteristics are the behavior's occurrence or nonoccurrence, its completion or incompletion, or its correctness or incorrectness" (p. 214). Binary data is the simplest and least time consuming system to utilize in the classroom. However, its greatest disadvantage is its insensitivity to behavior change. It is a useful tool for teachers who want to measure gross skill ability with new students to assess their current functioning level. Once such initial data has been collected however, teachers cannot rely heavily on them to measure pupil progress on more specific programs.

*Item Data, Levels of Dependence, and Counts Data.* A related measure is called "item data" (Gentry and Haring, 1976). Skills which have been analyzed (i.e. broken down into smaller, more discrete steps) can be checked "yes" or "no" according to the level of mastery (of an item) to find out where a student requires instruction to master an entire skill. The advantage of item data comes by providing specific information as to which components of a skill a child is having difficulty with. Item data does not, however, provide information on proficiency in skill acquisition, or an overall perspective of student progress.

Another related technique, "levels of dependence data," can be utilized if a child has little or no independent behavior in respect to the task or skill being programmed. Levels of dependence data would note the present level of dependence and movement toward more independent action. The following teaching sequence is utilized with respect to levels of dependence data:

1. The child is physically guided through the entire task. There can be no error.
2. Assistance is faded into a series of nudges, pushes or taps (physical assistance or prompts).
3. Verbal or gestural prompts are used.
4. The child independently performs the task without any special assistance or cues.

Levels of dependence data display would indicate the type of assistance the student needs to complete the task.

Another example of count data is "behavior counts." This type of data is expressed by recording a behavior each time it occurs. This differs from previous count systems in that the counts are taken for a fixed interval of time each day or session. During the school day (a fixed period of time) the number of times a student initiates a communicative interaction would be an appropriate behavior on which to obtain this type of "count data." Other examples would be number of bites taken at lunch or the number of steps correctly completed on a task sequence. Behavior count data is more sensitive in reflecting overall improvement than either "binary" or "item data" keeping, since a unit of time is involved.

*Trials to Criterion.* When a task is broken down into a series of small steps (task analyzed) and a student can perform each step with minimum assistance "trials to criterion data" may be utilized (Fredricks, Baldwin, Grove, Riggs, Furey, Moore, Jordan, Gage, Levak, Alrick, and Wadlow, 1976; Haring, 1977). The teacher selects an arbitrary desired performance level, such as three consecutive correct responses, then records the number of trials or opportunities needed to master each step before moving on to the next step. For instance, a teacher might require a student to correctly point out a classmate three consecutive times before moving on to the next step in the task of learning five other names in the room. As with checklists, trials to criterion require no manipulation of the data as do percent and frequency measures. Trials to criterion are simple and rely on count to show student progress and acquisition of targeted behaviors.

Gross and Cohen (1976) have identified the following assumptions, advantages and disadvantages of trials to criterion as a measure of behavior:

1) That the skill, behavior or task sequence is clearly and correctly established. The teacher must be certain that successive steps do in fact build upon one another and that achievement of one step or behavior assists in building the next behavior in the sequence or hierarchy.
2) That the number of trials designated by the teacher is appropriate to the task and to the learning characteristics of the child. Most trials to criterion measures involve a number of trials which are arbitrary. The teacher defines the total number of trials necessary, on the bases of his own preferences and beliefs. Who's to say that ten correct trials for three consecutive

days signify mastery of a task?

3) That the amount of time it takes to complete the trials is not important. Time is not a part of this measure. Although the behavioral criteria may be clearly specified in terms of movement and amount of assistance allowed, the amount of time it takes to perform a trial correctly does not matter (p. 33).

*Percent.* A system that is more sensitive to behavior changes in pupils is percent data. This system reflects the overall accuracy and quality with which the student completes the task correctly. Percent data represent the number of occurrences of a behavior compared to the number of possible occurrences of that behavior.

$$\text{Percent Correct} = \frac{\text{number of correct responses}}{\text{total number of responses (correct + error)}} \times 100$$

The quality and accuracy expressed in percent data shows the consistency with which a student is able to complete the step or task correctly. For example, a student took eight bites of food correctly without assistance and two bites of food incorrectly such that he required assistance. His percentage accuracy of the task would be 8/8+2 x 100 = 80%. As Haring (1977) explains "accuracy data expressed in percentages will help one to judge a child's consistency from one day to the next but will only be of use if the child is able to perform at least some behaviors without assistance" (p. 7).

A number of studies (Banerdt & Bricker, 1978; Fowler, Johnson, Whitman and Zukotynski, 1978; Jansen and Guess, 1978; Mithaug, 1978; Rusch, Greenwood, and Walker, 1978; Wehman, Renzaglia, Berry, Schutz and Karen, 1978) were searched to show typical data collected and measures used on severely handicapped students by teachers and researchers. The response units located in this search are shown on Table 1.

**Table 1**
**Response Units Found in Search**

| Unit | Type | Frequency |
|---|---|---|
| Count | No. of correct behaviors | 1 |
| | No. of behaviors | 4 |
| | No. of trials to criterion | 1 |
| Percent | % of responses correct | 25 |
| | % behaviors occurred | 5 |
| Frequency | Per Minute | 2 |
| | Per Day | 1 |

As in our classrooms, percent was the most frequently used measure found (e.g., Mithaug, 1978) in these studies. In the studies searched frequency was the second most used (e.g., Mithaug, 1978) and count, the simplest type of measure, third (e.g., Rusch, Greenwood, and Walker, 1978). Liberty (1976) found that on the units of time (x-axis or abscissa-horizontal line) consecutive sessions (or consecutive observation periods) were the most frequently employed measure used by teachers and researchers working with severely handicapped. Furthermore, in our rooms and in articles surveyed, consecutive days, sessions or periods were found to be the most frequently used measure of time, whether it was minutes, stimuli, trials, sessions by hour, data days, or probe sessions.

While percent is an easy measure to use and interpret, Gross and Cohen (1976) cite two disadvantages:

1) The use of a percent measure may impose an arbitrary ceiling on the amount of work, or number of responses a child is allowed to perform.
2) While percent tells us a great deal about accuracy, or the relationship of correct to error responses, it does not present a picture of proficiency. Proficiency is a picture of both accuracy and speed (p. 25).

*Rate/Frequency.* Proficiency may be measured by utilizing rate/frequency. Once a child has accurately acquired a skill, he must perform it quickly and easily to prove true mastery. If, for example, a child can put on his shirt without error, but it takes ten minutes or more to complete the task, it is not likely that the teacher or parents will consider the skill as being truly learned or functional. If speed as well as accuracy is important in a task and the teacher's goal is to increase performance speed, then rate or frequency data could be collected.

Rate, or frequency data are by far the most sensitive measures of behavior used in the classroom setting (Liberty, 1976). Rate/frequency data reflect accuracy of responding (as does percent) and also the fluency or speed of responding; endurance of responding; and the change across time. Rate/frequency data may, as well, be used to predict a pupil's performance in the future. Rate/frequency data are calculated by dividing the number of correct or incorrect behaviors by the number of minutes over which those behaviors were counted:

$$\text{Correct rate} = \frac{\text{\# of correct behaviors observed}}{\text{\# of minutes over which the behaviors were observed}}$$

For example, the rate of a student who matched eight picture cards correctly and two cards incorrectly during a two-minute observation period; would be: 8 picture cards divided by 2 minutes = 4.0 cards per minute, and 2 cards divided by 2 minutes = 1 card matched incorrectly. Utilizing the percent measurement mentioned previously this student performed the matching task with 80% accuracy. However, it may have taken one-half hour to perform those eight correct matching tasks, an important item of information about the child's behavior that is reflected only when using rate/frequency data. A behavior chart for recording rate/frequency data provided a visual display of the time required by the child to perform the behavior throughout acquisition.

Although rate/frequency data is highly sensitive to behavioral changes, a major disadvantage cited with rate/frequency measurement is that teachers are unfamiliar with the standard behavior chart paper (semi-log) that is utilized with this measure and experience difficulty using it since many procedures are rather complicated and mathematical (Liberty, 1976).

*Latency.* Another type of data involving time that is useful to teachers of the severely multiply handicapped is latency. Latency is defined as the amount of time it takes a student to respond after the teacher has given a cue to begin. The data is most easily obtained by using a stopwatch and timing the interval from the cue to the beginning of the response. Latency is most useful with classroom behaviors such as initiating a task upon receiving instructions (e.g., Sheri come!).

If the measure of latency of behavior is one time per day, the data is easily recorded and charted. However, if measures of a pupil's latency are taken throughout the school day data collection and charting become more complicated. In this case average latency is used:

$$\text{Average Latency} = \frac{\text{total amount of time between cues and responses}}{\text{total number of times measures were recorded (total number of trials)}}$$

For example, a student is told: "come to class," after arriving at school. Including recess, there are three times each day when the student enters the classroom. The teacher, in this instance, is interested in decreasing the amount of time it takes the pupil to enter the classroom after being called. During the intervention period the teacher calculates the average time this student takes in entering the classroom to determine if there is a decrease in the amount of time required to enter the classroom upon cue.

A latency measure is sensitive to the amount of time required for a response to begin, but it does not measure other behavioral dimensions such as quality and quantity of responses. (e.g., the pupil rolled, crawled, walked, or fell 15 times before arriving).

*Duration.* Duration data reflect changes in the length of time of a particular behavior. The data is collected by noting the amount of time which passes from the beginning of the behavior to the end of the behavior. Examples of duration data include: time head is up, time a student is sitting or standing, and length of on-task behavior.

As with latency, duration data only provides time information about behavior, with no indication of the quantity or the quality of the behavior. If taken once a day, the data is easily recorded and charted; otherwise, averages, as with duration may be utilized. For example, a teacher wishes to increase the time a student wears his glasses or sits by himself. Along with an intervention program on a particular behavior, the duration of this behavior is recorded. Duration, in this case, is an excellent measure of behavioral change. Duration data might also be useful when trying to measure endurance on vocational or pre-vocational tasks. As pointed out with every measurement system employed, a teacher needs to carefully examine the most critical aspects of the behaviors on which data is to be collected. The information sought must be reflected in the data that is collected.

### Selecting a Measurement System and Data Collection Practicality

A teacher must choose a measurement technique which accurately reflects pupil progress. In doing so, the teacher must consider certain factors related to data collection. Firstly, a teacher must be familiar with the collection and use of the different types of data described before implementation. Secondly, the teacher should be conscious of data practicality itself, considering whether the data collection process, in itself, would be important enough to compensate for interrupting the class instruction. The teacher must also decide whether the data will be immediately displayed or charted later in the day. As stated previously, data must be in some way comparable on a day to day or session to session basis in order to be accurate. Data should be taken frequently and behaviors defined in small steps to reflect accuracy, proficiency, endurance, latency, or duration.

Therefore, instructional procedure and general classroom management; teaching strategy and tactics, time, assistance in the classroom; and equipment available for counting and timing are very important factors for a teacher to consider in implementing data collection in the classroom. Finally, data needs to be relevant to specific programs in the classroom. The more sensitive a measurement unit is in monitoring progress (or lack of it), the more effective the feedback the teacher will gain from it.

In selecting a measurement technique the teacher must remain cognizant to the fact that the most salient dimensions of a behavior may change over time and thus a change in the type of data collected on a particular behavior is needed. Once the behavior and its critical dimensions are defined, a decision about the type of data to be collected will be simplified.

### Rules for Decision Making

As White and Liberty (1974) suggest, there are no magic rules or numbers to decide when a student has acquired a new skill fluently. When advancing severely multiply handicapped students to new steps or tasks Gentry and Haring (1976) recommend that teachers:

1) determine a performance criteria representing acquisition, mastery, or proficiency;
2) determine how many times, or for how long a student must achieve a criterion of performance before advancing (e.g. number of tasks or steps, or how many trials before advancement);
3) select a rate of progress before intervention begins and make changes as needed;
4) alter teaching procedures if a student is not making performance gains;
5) avoid changing teaching procedures when current procedures are successful.

Horner, et al. (1976) discuss the importance of baselining behaviors before decisions are made. Baseline refers to the process of taking data on specified behaviors before actual training begins. With an ascending baseline the student is learning the behavior without training. With a decending baseline, they suggest beginning training immediately. With a stable baseline below criterion level, training could also begin immediately. When baseline is at or above criterion level, a new task should be selected. The opposite holds for decreasing behaviors.

Horner, et al. (1976) further suggest that a student reach 80-90% criterion on acquisition of a new skill for a specific number of days. If an abnormally long period of time is required to complete the task a different measurement system (e.g. duration) could be selected.

### Conclusion

Whatever type of response unit is measured, it is of primary importance to select a display system that will convey information to the teacher as to whether or not intervention programs are successful, require alteration, or are maintaining or improving student performance (Haring, 1977; Liberty, 1976; White and Liberty, 1976).

Teachers need to determine not only successful intervention procedures, but also measure procedures which provide the most information with the least expenditure of teacher time and energy in successful student programming. Each data collection procedure presented has been proven effective for certain types of behaviors and programs for severely multiply handicapped.

Liberty (1976) reminds teachers that merely collecting appropriate data in and of itself will not guarantee success of instructional programs. "It will, however, guarantee that information necessary to appropriate decisions regarding change or continuance of the program is available. Without such information, teachers will proceed in the dark, turning on lights without judging whether or not they illuminate the way for the students they lead" (p. 27).

### References

A complete list of references will be supplied upon request.

*Marty Mattison II is Teacher of Severely Multiply Handicapped children, Kainalu Elementary School. He received an M.Ed. in Special Education from the University of Hawaii.*

*Rebecca Rosenberg is Teacher of Severely Multiply Handicapped children, Honowai Elementary School. She received an M.Ed. in Special Education from the University of Hawaii. Both Mr. Mattison and Ms. Rosenberg are among co-authors of* Holo I Mua: A guide for teachers of severely multiply handicapped, *curriculum guide for SMH teachers, State of Hawaii.*