
Standard 9 Data Collection and Evaluation



Standard 9 requires the student:

- A. Understand different data recording procedures.
- B. Understand how to graph data.
- C. Understand how to analyze data and plan for interventions.

Key Terms for Standard 9

Baseline data: Test data taken on a skill before a teaching program is initiated.

Probe data: Data taken on a skill under testing conditions once a teaching program has been initiated.

A. Understand different data recording procedures.

To evaluate the impact of providing therapy to a child, the DS must formulate specific strategies for measurement.

Three basic reasons for developing strategies are:

- (a) To document what has occurred,
- (b) To identify the variables responsible for the occurrence, and
- (c) To enable the DS to better predict future performance, because performance helps the DS decide if changes to the program are necessary.

Continuous evaluation of a child's progress is an important part of the teaching-learning process and should result in the identification of concerns and difficulties and provide guidance in revising therapy decisions (Cullen & Pratt, 1992).

Sometimes methods of collecting objective data are so complex that the DS, paraprofessional, and other team members cannot easily determine how to collect the data. In addition, time to collect data may be minimal depending on the student's support needs and the situation. As a result, data collection procedures must be very clear, simple, convenient, and planned. A child's IPP/IEP goals and objectives must be clearly written and easily measured. When objectives are vaguely written, collecting data is much more difficult. Clearly written objectives should specify exactly what is being measured, under what conditions, and the level of mastery expected of the child.

Example 1

Vague IPP/IEP Objective: Susie will improve communication skills 80% of the time.

Specific IPP/IEP Objective: When a peer asks her a question, Susie will look at him or her and then point to the appropriate person, object, or picture on her communication device within 3 seconds for 8 of 10 consecutive questions.

In the vague example, what the word "improve" refers to is not clear. The "communication skills" being addressed are also not clear. Each person working with Susie may have a different idea of what should be measured. Consistency across service providers in gathering meaningful data may be lost without a clearer explanation of the objective. The criterion of "80% of the time" is confusing because it is not clear what time period is being considered. It would be unusual for someone to observe a child's opportunities for communication throughout the school year.

In the more specific example, the situation/conditions are stated (“when a peer asks Susie a question”) she is required to respond to the question by looking and pointing (Susie’s required behavior). It also states how quickly and how often she must respond to reach criterion. It is important to note that before a team can develop the criterion for a goal or objective they must know Susie’s present level of performance.

Baseline data is test data that is collected before a teaching program is initiated. During this phase the child is allowed to perform a task independently without feedback, reinforcement, or prompting. Initial data collection is referred to as baseline and should be conducted at least once before the initiation of therapy (Brown & Snell, 2000). However, baseline data should be collected for at least five days before implementing specific behavioral interventions (Alberto & Troutman, 2003).

Baseline data creates a record of whether a child performs the step in a task analysis correctly (+) or incorrectly (-). The number correct can be recorded as a percentage.

Number of steps performed correctly/total step x 100 = % of steps performed correctly.

An ecological inventory may serve as the source of a child’s baseline data depending upon the skill or activity being assessed. A DS might want to compare baseline and instructional data on a graph. A graph can create a visual representation of the child’s progress over time.

Once baseline data has been collected the DS can use it to compare it to data collected during the **instructional phase or intervention phase** (of specific inappropriate behaviors). The instructional phase of a child’s training program will require different data collection methods although the DS may conduct a **probe assessment** during training to compare the child’s performance to the baseline data. A probe would be conducted at the beginning of a training session and would not provide feedback, reinforcement, or prompts to the child.

There are several ways to measure a child’s performance during training. The strategy selected should match the behavior to be measured and the situation.

Frequency Recording

Frequency recording, sometimes referred to as event recording, measures the number of times a behavior occurs within a period of time (e.g., number of times child initiates during therapy, number of times a child hits himself during a certain period of time). This kind of data collection method is usually used when the objective is to increase or decrease the number of times a child engages in a certain behavior.

To have meaningful frequency data, it is necessary that comparisons be made only if training sessions are the same length of time. For example, the DS might note that a child has hit himself 10 times on Monday and 3 times on Thursday. This may look like the child has decreased the

behavior. However, if the DS observed for an hour on Monday and 15 minutes on Thursday it is not possible to conclude that the child decreased the behavior.

Behaviors that are collected in this manner should be discrete and have a clear beginning and end (Alberto & Troutman, 2003). Frequency data should not be collected on:

- 1) Behaviors that occur at a high frequency that may reflect an inaccurate count (e.g., flapping, rocking).
- 2) Behaviors that can occur for an extended period of time (e.g., thumbsucking).

Frequency recording is relatively easy to do. A notation can be made on an index card or piece of paper. At the conclusion of a training session, the information can be tallied and transferred to a formal data sheet.

Percentage

A percentage is used to compare the number of correct responses with the number of opportunities provided to the child (e.g., child performed 6 out of 12 steps of an activity). It is not appropriate to use a percentage when there is no ceiling on the number of opportunities a student has to perform the skill/behavior. Percentage is calculated by dividing the number of behaviors observed by the number of opportunities the child has to perform a behavior.

“Jamie will greet her peers 80% of the time” does not provide a ceiling on how many opportunities Jamie had to greet her friends.

Using a percentage score is also useful when the number of opportunities varies across training sessions. For example, if a child has five opportunities to greet their peers and only greets them twice during a period of time, the child will have greeted their peers 40% of the time.

Rate

Rate can be used to determine the frequency of a behavior and its relation to time. The score is expressed as a ratio of the number of behaviors divided by the amount of time. Rate data can also reflect the accuracy, speed, or fluency of a response.

Henry threw his spoon on the floor 15 times in a 30-minute session, or 0.5 times per minute.

Jane went from washing 5 dishes in 30 minutes, or .17 per minute, to 10 dishes in 30 minutes, or .33 per minute.

This type of measurement is useful when the child's number of opportunities to perform a behavior varies across training sessions; for example, to measure the number of spoonfuls of food a child eats when the length of snack time varies. It can also reflect a child's proficiency on a task.

Duration

This type of measurement is used if there is concern about the amount of time a child is engaged in a specific behavior or activity. Sometimes, it is necessary for the child to increase the time they are engaged in an activity (e.g., brushing teeth), and sometimes it is desirable for a child to decrease the amount of time in an activity (e.g., self-injurious behavior). Duration measures the total time a behavior occurs.

Duration recording, like frequency recording, can be used for discrete behaviors – those with a clear beginning and end. This type of measurement can be recorded in three ways:

(a) total duration, (b) a percentage of time, and (c) by measuring each occurrence. (Brown & Snell, 2000).

Total Duration and Percentage of Time

This measures how long a student is engaged in a behavior within a limited period of time. This activity may or may not be continuous. The behavior must be clearly defined. For example, a DS might be interested in collecting data relating to a child's "on task" behavior. The DS would first clearly define the on-task behavior to make sure that it is possible to clearly determine the onset and termination of the behavior. The DS could then monitor the behavior using a stopwatch. As soon as the child stops being "on-task" (e.g., participates in a self-stimulatory behavior), the DS would stop the stopwatch. The DS would start the stopwatch if and when the child began to engage again in the task. The amount of time accumulated on the stopwatch at the end of the allotted task time would reflect the total duration.

$$\frac{\text{Total duration of behavior}}{\text{Length of observation period}} = \frac{5 \text{ min}}{15 \text{ min}} = 33\%$$

Measuring each occurrence

This kind of measurement is simple and can be accurate if the behavior is clearly defined. It also can provide the DS with the frequency of each occurrence. For example, the DS might know the percentage of time a child was "on-task" (33%); however, this figure does not indicate if the child was on task for the entire 5 minutes in a row or 45 seconds and kept returning to the task. This kind of information can be important for determining the type of intervention to use for the

child to increase the on-task behavior. Measuring each occurrence will give the DS this kind of information although it is more time consuming than simply measuring the total duration and percentage of time the child is engaged in the task.

To measure each occurrence, the DS would start the stopwatch when the child began to engage in the task, turn the stopwatch off when they stopped the task, and then record the duration on a data sheet. The DS would then return the stopwatch to 0. When the child began the task again, the DS would start the stopwatch and keep it going until the child stopped the task. When the child stopped again, the DS would record the duration, and so on. At the end of the observation period (e.g., 15 minutes) the DS would have a record of total duration (e.g., 5 minutes), as well as a count of the number of times the child started and stopped the task (e.g., 8 times). The goal would perhaps be to have the child increase the duration of time spent on the task and to decrease the number of times the child was distracted from working.

The advantage of this type of measurement is that it gives the DS a record of the length of the occurrence of a behavior and can also record the duration of each “incident” of a behavior. It does require the DS to be very attentive during a training session. A stopwatch is also necessary to capture accurate data. This kind of data collection method would not be appropriate for frequent behaviors that last a short period of time (e.g., tantrum).

Interval Recording

Interval recording is useful for those behaviors that do not have specific start or stop times and may vary in length (e.g., stereotypic or aggressive behavior) (Alberto & Troutman, 2003). Instead of counting each time a behavior occurs the DS counts the number of intervals of time within a sample observation period in which the behavior has occurred. The DS defines a certain period of time in which the behavior will be observed. This period is then divided into equal intervals. Typically, the intervals are no longer than 30 seconds (Alberto & Troutman, 2003). The series of boxes are drawn to represent the different intervals. The DS simply notes whether the behavior occurred (+) or did not occur (-) at any time during the interval. Each box has one mark in it. The actual number of occurrences is not noted within an interval.

There are two kinds of interval recording: **whole interval** and **partial interval**. In whole interval recording, the DS records if a behavior occurred continuously throughout the interval (e.g., being off task, screaming).

In partial interval recording, the DS documents if the behavior occurs at any time during the interval, rather than if it occurs continuously throughout the interval. Once a behavior is noted, it is not necessary to observe the behavior for the rest of the interval. The number of times the behavior occurs is not noted. Interval recording provides an estimate of the occurrence of the behavior, therefore, only limited conclusions can be drawn from the data (Alberto & Troutman, 2003).

The selection of the appropriate interval recording method should be guided by the child's behavior and the goals of the intervention. If the DS is trying to increase a behavior that should occur on a continuous basis, such as staying on task, then the whole interval recording is appropriate. But, if the behavior occurs for a shorter period of time, such as making eye contact during interactions, then a partial recording method would be more appropriate.

Sometimes a third person is required to take interval recording data. It is difficult to provide therapy and collect interval data simultaneously. The DS must interact with the child, observe the specific time intervals, and note the occurrence or non-occurrence of the behavior. Even though the DS might use a stopwatch or watch with a second hand, checking the interval time can cause the DS to miss a behavior. Sometimes an observer will build in a short time period (e.g., 5 seconds) between observations. Longer intervals (e.g., 3-5 minutes), may require the use of a recording with timed beeps that would alert the DS to the interval (Alberto & Troutman, 2003). The DS should always be sensitive to the environment and as unobtrusive as possible.

When the goal of intervention is to reduce a behavior it is best to use the method that provides the most rigorous data. For example, if the whole interval recording method is used to reduce a child's self-injurious behavior, then the DS would check only if the behavior occurred **throughout** the interval. Using this approach would not record any incident of self-injury unless it had been continuous. Therefore, the partial interval method would be more appropriate.

Whole interval recording is useful when it is important to know that a behavior is not interrupted.

Partial interval recording is useful when it is important to document behaviors that may only occur for a brief moment.

Time Samples

Time sampling is another type of interval recording that can be used in a training session. As with whole and partial interval recording a specified observation time (e.g. 15 minutes) is divided into small periods (e.g., 5-minute intervals). Instead of observing the child's behavior throughout the interval, the DS only observes the child at the end of the interval. This kind of method typically uses longer intervals (minutes) than does interval recording (seconds). The DS records whether or not the child was engaging in the behavior at the end of each interval.

Time sampling is more flexible than interval recording in that the DS can set up random intervals within an observation period. For example, the DS might decide to set up the observation period for an hour and select five random times to observe in the course of the hour.

This method also allows the DS to provide therapy and collect data at the same time. But, like interval recording, time sampling only provides an estimate of the targeted behavior. Time sampling is most appropriate for measuring behaviors that are fairly frequent and that occur for long periods of time.

For low-frequency and short-duration behaviors, time sampling is less accurate than interval recording. The less frequent or briefer the behavior, the shorter the interval must be. As the interval gets longer, the similarity between the data recorded and the actual occurrences is likely to decrease (Alberto & Troutman, 2003).

Interobserver Reliability

Interobserver reliability is assessed to determine whether the child's behavior is being recorded accurately. One way to ensure that the data is accurate is to have two independent observers record the behavior at the same time then compare the two observations. The percent of the interobserver reliability can be calculated by dividing the number of agreements between the two observers by the number of agreements plus disagreements and multiplying by 100. The result is called a percent of agreement. Generally, a reliability coefficient of .80, or 80% is considered acceptable (Alberto & Troutman, 2003).

Agreements

$\times 100 = \text{Percentage of Agreement}$

Agreements + Disagreements

B. Understand how to graph data.

Graphing Data

A graph is made up of two axes. The abscissa, or the x-axis, is the horizontal line. The abscissa usually represents the time frame of a measurement (e.g., a data point would reflect data from a day, week or month). The ordinate, or the y-axis, is the vertical line. It represents the behavior being measured (e.g. time on task, hitting) and the measurement that was used (e.g., duration, percent).

Data can be more effectively interpreted if it is converted into a graph. A graph will allow the DS and the child's family to have a visual representation of the child's progress.

Data must be converted into a single numeric form before it can be graphed. Different data methods will provide different kinds of scores (e.g., rate, percentage, count). The type of data conversion method will depend upon the type of collection method used by the DS and the kind of data reported.

The following chart is a summary of data conversion procedures:

Recording Method	Conversion	
1) Permanent product	Number of occurrences	If both time/opportunities are to respond are constant
2) Frequency recording	Percentage	If time is constant but opportunities vary
	Report rate	If both time and opportunities vary, of it time varies and opportunities are constant
3) Interval recording	Report #	If constant
4) Time sampling	Report percentage of intervals	During or at the end of behavior
5) Duration	Report # of seconds minutes/hours	For which the behavior occurred



Student Note: Remember to:

- 1) Label the ordinate or vertical line (y-axis) with the behavior being measured and type of measurement used (number of verbalizations during lunch).
- 2) Next, divide the ordinate into equal intervals that cover the possible range of data (e.g., 0-100%, 0-50). If there is no range then it should be divided from the baseline level to the target level leaving some room for variability.
- 3) Data points should be connected by a straight line but not connected across phases (e.g., baseline, intervention).
- 4) Date the graph along the horizontal line (x-axis) using time intervals that the data was collected (e.g., daily, weekly, monthly).
- 5) When analyzing data take into account any missed sessions.

Self-Graphing Data Sheets

The DS may find it helpful to combine the data recording sheet and the graph. This has been done to measure prompting levels, task sequences, and task analysis (Brown & Snell, 2000).

This kind of data sheet can be used to record probe data by making a slash (/) through the step if the child responds independently and an X through the step number if the child needs assistance to complete the step. At the conclusion of the training session, the DS would add the number of slashes and then circle the number correct for the day. A graph is formed by connecting the circles across days.

C. Understand how to analyze data and plan for interventions.

Once data is collected it must be analyzed and interpreted to ensure that a child is making progress. The child's team must recognize the behavior to be changed, the **dependent variable**, like a child's on task behavior or hitting. The strategy used in changing the behavior is the **independent variable**, such as different kind of prompting procedures, reinforcement, environmental factors, etc. Changes in the dependent variable "depend" on changes in the independent variable.

The team must be aware of when a change needs to occur. "When a student's progress has been below aim for approximately three out of five sessions or if the trend is flat or descending (or ascending for deceleration programs), teachers should attempt to determine why and make program improvements based on their analysis of the data" (Brown & Snell, 2000, p. 201).

The team needs to review several sources of data as they plan changes to a child's program.

These sources include:

- 1) **Anecdotal records** – notes and comments from other staff or family members about the behavior or performance of the child (e.g., illness, changes in staff).
- 2) **Ungraphed training data** – information that might not appear on the graph but provides the DS with response-by-response information about a student.
- 3) **Ungraphed probe data** – what steps a child can do on a targeted activity or routine without prompts or feedback.
- 4) **Graphed training data** on targeted skills.
- 5) **Graphed test data** on targeted skills.

By using a variety of data sources the team can develop explanations for the child's lack of progress and implement program changes. No matter what system of data collection and analysis is used, it is important that the child's team use the data they collect to monitor progress and improve the program.

Points to Remember

- ✓ Data collection procedures must be clear, simple, convenient, and planned.
- ✓ Use multiple pieces of data to evaluate a student's performance/behavior.
- ✓ Use data constructively.



Student Project: Review ~~four~~ different IPP's and discuss the data collection methods used for various objectives. Please indicate how and why you would make changes in any of these methods, or why you would not make changes.

Use one type of data collection and do a probe on your client. Write an objective using your gathered data.

References

- Alberto, P.A., & Troutman, A.C. (2003). *Applied behavior analysis for teachers*, (6th ed). Upper Saddle River, NJ: Prentice Hall.
- Brown, F., & Snell, M.E. (2000). *Measurement, Analysis, and Evaluation*. In M.E. Snell & F. Brown (Eds.). *Instruction of students with severe disabilities* (5th ed.), (pp. 173-206). Upper Saddle Rive, NJ: Merrill/Prentice Hall.
- Cullen, B., & Pratt, T. (1992). *Measuring and reporting student progress*. In S. Stainback & W. Stainback (Eds.). *Curriculum considerations in inclusive classrooms: Facilitating learning for all students*, (pp.175-196). Baltimore: Paul H. Brookes Publishing Co.

Ways to Measure a Child's Performance

- Frequency Recording
- Percentage
- Rate
- Duration
- Interval Recording
- Time Samples

Event Recording Data Sheet

Student: _____

Observer: _____

Behavior _____

	Time		Notaton of occurences	Total occurrences
	Start	Stop		