A Direct Observation Approach to Measuring Classroom Behavior: Procedures and Application.


Bureau of Education for the Handicapped (DHEW/ OE), Washington, D.C.

IPLDRP-6.

Apr 79

300-77-0491

32p.: See also EC 123 901-925.

Behavior Problems; Case Studies; Classroom Observation Techniques; Elementary Education; Evaluation Methods; Informal Assessment; Measurement Techniques; Peer Groups; Student Evaluation

An approach for systematically observing and recording behavioral difference in the classroom is described, and its use with 11 behavior problem elementary students is discussed. The author points out the importance of peer sampling and the use of multiple observations to avoid erroneous conclusions. A case study of a second grade child illustrates the potential use of the observation system. (CI)
A DIRECT OBSERVATION APPROACH TO MEASURING CLASSROOM BEHAVIOR: PROCEDURES AND APPLICATION

Stanley L. Deno
The Institute for Research on Learning Disabilities is supported by a contract (300-77-0491) with the Bureau of Education for the Handicapped, Department of Health, Education, and Welfare, U.S. Office of Education, through Title VI-G of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled children. Research activities are organized into eight major areas:

I. Adequacy of Norm-Referenced Data for Prediction of Success

II. Computer Simulation Research on the Assessment/Decision-making/Intervention Process

III. Comparative Research on Children Labeled LD and Children Failing Academically but not Labeled LD

IV. Surveys on In-the-Field Assessment, Decision Making, and Intervention

V. Ethological Research on Placement Team Decision Making

VI. Bias Following Assessment

VII. Reliability and Validity of Formative Evaluation Procedures

VIII. Data-Utilization Systems in Instructional Programming

Additional information on these research areas may be obtained by writing to the Editor at the Institute.

The research reported herein was conducted under government sponsorship. Contractors are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official position of the Bureau of Education for the Handicapped.
Research Report No. 6

A DIRECT OBSERVATION APPROACH TO MEASURING CLASSROOM BEHAVIOR:
PROCEDURES AND APPLICATION

Stanley L. Deno
Institute for Research on Learning Disabilities
University of Minnesota

April, 1979
Abstract

As a result of the countervailing pressures to both effectively respond to classroom disorder and yet protect student rights, principals, teachers, special educators, and school psychologists have become sensitive to the need for objectively documenting the basis for interventions. While evidence has accumulated that behavioral ratings are susceptible to the biasing effects of pupil characteristics and prior labeling, other evidence can be found that direct observation and recording are not so influenced. Presented here is a procedure for systematically observing and recording behavioral difference; it has been used successfully by special educators over a period of five years to document the existence of behavioral difference and to evaluate intervention effectiveness.
A Direct Observation Approach to Measuring Classroom Behavior: Procedures and Application

Virtually every teacher has at least one student who might variously be described as "a behavior problem," "inattentive," "hyperactive," "acting out," or "disruptive." The problems of managing social behavior in the classroom are so pervasive, in fact, that "lack of discipline" has been identified by the American public in eight of the last nine years (Gallup, 1978) as the major problem facing the schools. Disorderly classroom behavior is so common that it has become normal for a child to be identified at least once as a behavior problem during the elementary school years (Rubin & Balow, 1978).

At the same time, the courts have acted clearly to establish the basic rights of students under the Constitution. Most recently, lawmakers have made it clear to educators that despite the ease with which they might identify children as socially deviant, they must not selectively segregate a child without ensuring that the child's rights to due process of law have been met (Mills v. Board of Education, 1972). Further, if the school's recommendation is to make substantive revisions in the child's environment, those revisions must not only include due process protections, but their effects must also be carefully monitored. As a result of the countervailing pressures to both effectively respond to classroom disorder and yet to protect the student's rights, principals, teachers, special educators, and school psychologists have become sensitive to the need to objectively document the basis for interventions in children's lives. How, then, shall this be done?
One approach has been to use behavior checklists or rating scales completed by the classroom teacher or someone who has observed the student in question. When such instruments are used, however, the probability of bias is significant (cf. Foster, Ysseldyke, & Reese, 1975; Ysseldyke & Foster, 1978). A preferable alternative seems to be the use of direct observation and recording of behavior which is less susceptible to biasing (Madle, Neisworth, & Kurtz, 1979), and, consequently, better meets the requirement that behavioral difference be objectively documented.

Since 1972, the author has been involved in training teachers to intervene in their role as special education resource teachers. In their role as intervention managers the teachers have learned to assess difference in classroom behavior using a direct observation system developed in connection with a Special Project funded by the Bureau of Education for the Handicapped (USOE).

The observation system is based on the assumption that any label applied to a child which identifies that child as a conduct problem implies that the child's behavior differs significantly from that of his or her peers. Proceeding from that assumption, we have used the observation system to objectively determine the existence, and the extent, of that implied behavioral difference. The remainder of this paper is devoted to describing the observation system used, to providing some data on its validity, and to illustrating its use.¹

The Observation System

The recording procedures are only briefly described here to provide the reader with a general understanding of the observation system. A
more complete description is contained in a manual which may be obtained upon request from the author.

What to Record

The first concern in an observation and recording system is with what behaviors shall be observed and recorded. While in some approaches the decision regarding target behaviors must await the individual case, the position taken here is that a set of behaviors can be identified which fairly represents the categories of concern for most classroom teachers. The categories are "noise," "out of place," "physical contact," and "off task." The number of categories is small so that the recorded behavior is likely to be high enough to warrant attention, and so that results are relatively easy to interpret by all concerned parties.

Definitions. To develop consistency among observers, the following standard definitions of the categories of behavior are used:

1. **Noise** is defined as any sounds created by the child which distract either another student (or students) or the teacher from the business at hand. The noise may be generated vocally (and includes "talk outs" or unintelligible sounds) or non-vocally (as "tapping a pencil" or "snapping fingers").

2. **Out of place** is defined as any movement beyond the either explicitly or implicitly defined boundaries in which the child is allowed movement. If the child is seated at his desk, then movement of any sort out of the seat is "out of place." If the child is working with a group, then leaving the group is "out of place."

3. **Physical contact or destruction** is defined as any contact with another person or another person's property which is unacceptable to that other person. Kicking, hitting, pushing, tearing, breaking, taking, are categorized as physical contact or destruction.

4. **Off task** is defined as any movement off of a prescribed activity which does not fall into one of the three previously defined categories. "Looking around," "staring into space," "doodling," or any observable movement off of the task at hand is included.
5. **Other.** While the behaviors defined above serve as a reasonable basis for most observations, individual cases may arise where other behaviors should be recorded. Children may be identified who do not communicate or who do not interact. In such instances either "self-initiated utterances" or "self-initiated contacts" may be added, defined, and recorded. Generally, however, the first four categories will encompass many of the discrete categories which might be considered, and the "other" category should only be used if absolutely necessary to clarify the "problem" identified by the teacher.

**Frequency Records**

The rationale for developing a system which focuses on what some would describe as exclusively negative behaviors should be given. First, to observe the reciprocal of each behavior (i.e., "quiet," "in place," "no physical contact," "on task") requires use of a recording system which describes no action on the part of the student. Recording procedures which describe no action, such as interval recording, time sampling recording, or duration recording, generally require more observer training and technology than is desirable if the system is to be widely used. Second, to record the non-occurrence of behavior is to describe what has been referred to elsewhere as "the statue" (White & Haring, 1976) and to focus attention upon the student who will "sit, be quiet, and pay attention" -- a value position which has been criticized elsewhere (Winett & Winkler, 1972). Finally, and perhaps most importantly, the purpose of the recording system is to aid in determining whether the frequency of the so called negative behaviors (which are, presumably, the basis for labeling the child with one of the common misconduct labels) is, in fact, excessive. To not obtain a record of those behaviors would be a failure to directly measure the presumed basis for the label.
Interobserver agreement. As is true for any recording system, the definitions of the behaviors alone cannot be used as the exclusive basis for ensuring consistency among observers; agreement among observers is a necessary condition for using the system. For that reason, any time observations are being made, on at least one occasion, two or more people observe and record the same behaviors at the same time. The consistency between these observations should then be determined by computing a coefficient of interobserver agreement. If two people cannot reasonably agree on how often something is happening, then their disagreement should be resolved before further attempts to record the child's behavior are made.

Procedures. Figure 1 presents a recording sheet to be used when observing behavior. Each box represents an interval of one minute. The first box in each row used for the target child's behavior represents the first minute during which that child's behavior is to be recorded. That is, all of the behaviors to be recorded for the target child are recorded simultaneously during the first minute. The recording method used has been variously referred to as "frequency," "event," or "ally" recording (cf. Hall, 1971; or Gelfand & Hartman, 1975). Each time any one of the behaviors occurs during the first observation minute, the recorder continues to place tally marks in the first box for each occurrence of the behavior until that minute has passed.

At the end of the first minute of recording, the observer turns attention away from the target child to another child in the classroom.
or group. During this second minute of observation, the behavior of this child (called peer #1) is observed and recorded in the same manner as was just completed with the target child. Tally marks for peer #1 are entered in the first boxes for the peer's behavior until the second minute of observation has been completed. The cycles of recording target child behavior for a minute and then another peer's behavior occur throughout the recording period.

The peer sampling procedures described here were derived from a coding system for observing classroom behavior developed by Cobb and Ray (1970). The system was modified subsequently and used by Patterson, Cobb, and Ray (1972) to contrast the behavior of a targeted student with that of his or her peers. More recently, Walker and Hops (1976) presented data on the use of normative peer sampling as a standard for evaluating treatment effects. They concluded that peer sampling is a useful procedure not only in research, but also in evaluating intervention effects because it provided a framework for interpreting daily fluctuations in behavior and for estimating the level of behavior necessary to reintegrate a student with peers.

Within the perspective presented here, the primary purpose for observing peer behavior is to assess the basis for the teachers' descriptions of the target child's behavior. The assessment is accomplished by obtaining a precise measure of the relative difference between the frequency of target child behavior on the one hand, and the frequency of peer behavior on the other. A comparison such as this is not useful unless two conditions are met:
1. The peers constitute a group in which the target child is actually expected to hold membership (and, therefore, his or her behavior in that group is relevant).

2. The peer behavior which is recorded is actually a representative sample of the behavior of the total peer group.

The first condition is usually met if the child is participating in a regular classroom in a conventionally organized school program. In these cases, the classmates are the reference group. In school programs where groups shift frequently (as in modular schedules, ungraded programs, or open schools), more care must be taken to determine the appropriate reference group. In all cases, the teacher who has identified the child should be drawn into identifying the appropriate peer group for the target child since it is the teacher's expectations which are at the basis of the initial referral.

The second condition can be met only after the appropriate peer group has been identified, and is met by systematically obtaining representativeness. Prior to beginning recording, the observer should determine the process by which the peers will be selected during the observation. The process used should be one which eliminates the need for arbitrary decisions on the part of the observer since there are almost certain to be biased by unpredicted situational factors (e.g., where the target child is sitting, or who happens to be making noise at the moment). It is important to point out in this context that the observer is not trying to assess the target child's behavior relative to other children "most like him or her" (i.e., one does not attempt to select other children who are likely to "make a lot of noise" or be "out of place" a lot).
Summarizing results. Once the data have been obtained it becomes important to summarize results in a form readily comprehensible to teacher, parent, administrator, and child. No single summary procedure will always be satisfactory. The summary procedure recommended here, however, is called the "discrepancy ratio." It is the relative difference between the median level of the target child's behavior and the median level of the peers' behavior. The child who is off task four times per minute while his or her peers are off task two times per minute is a child who would be described as "off task two times (2X) more than his or her peers." Illustrations of this approach to summarizing the difference in the social behavior of a child and his or her peers are presented in the data which follow.

Validity and Utility of the System

Validity of measurement in behavioral recording is generally not an issue since the behavior to be observed is specified by the person concerned with that behavior and measurement may involve simply tallying each occurrence. Content validity is thereby ensured; concurrent, predictive, and construct validity are not an issue since inferences beyond the data are not made.

In the case of observational systems which use standard behavior categories such as those presented here, however, the person concerned with the behavior (typically the teacher) has not selected or specified the target behaviors. Further, the inference may eventually be made that the target child is, in some sense, socially deviant. To provide a modicum of data on the validity, and the sensitivity, of the observation system described here a study was conducted to determine whether
students identified by their teachers as socially deviant did, in fact, behave differently from their peers when the present observation system was applied.

Method

The study was conducted in a midwestern inner city elementary school with an enrollment of approximately 550 pupils. The school is located intermediate to schools with the highest and lowest income families, and its minority enrollment at the time of the study was about ten percent.

Observation sample. Eleven students were selected for observation by simply asking the teachers to "identify the child in your class who is having the most difficulty adjusting socially." The sample of target students included at least one child from every grade from K-6, with a sex distribution of six boys and five girls. Peers were selected by the observer using a systematic sampling strategy to avoid bias. Ten different peers were observed on each observation.

Procedure. Observations were conducted during periods of structured academic work (i.e., independent "seat work" during reading, math, etc.). The observation periods ranged from 8:30 a.m. to 1:30 p.m. and each was twenty minutes long, providing ten minutes of data on a target child, and ten minutes of data on the peer sample. Five, not necessarily consecutive, days of observations were made for each target child and peer sample.

Results

The results for all observations are summarized in Table 1.

---

Insert Table 1 about here

---
As Table 1 illustrates, the greatest relative difference between target students and their peers existed in the behavioral categories called "out of place" and "off task"; more target students differed by a factor of at least two on those categories. As might be expected in a structured academic period, "physical contact" or "destruction" occurred with the least frequency. A room-by-room analysis of the data revealed considerable variation in mean frequencies of the peer sample for a given behavioral category. Off task behavior, for example, ranged from a low of .04 occurrences per minute in three rooms to .11 occurrences per minute in a different three rooms.

Analysis of individual target student data revealed that each student differed from his or her peers on at least one of the behavioral categories by a factor of two or more times. (That is to say, each student selected as "deviant" by the teachers performed at least one of the four behaviors twice as much as his or her peer sample.) Figures 2, 3, and 4 present the individual data for three different target students and their respective peer samples.

The individual data are presented to illustrate the different configurations which the observation system might reveal and the variability in day-to-day behavior characteristic of both target students and peers. As may be seen, an observed difference in median frequency of occurrence between the target child and the peer sample was obtained for Pa and Pe on the category called "out of place," but not obtained for Oa. A difference was obtained in "off task" for Pa and Ma, but not for Pe. In Pe's case, however, the subject was less noisy than the subject's peers.
None of the three target students differed in median frequency of occurrence for the category "physical contact"; however, a difference would clearly be obtained for Ma if mean rather than median frequency of occurrence was used as the measure of central tendency. It is worth further mention that the median difference obtained for "physical contact/destruction" between target students and peers was smaller than the median differences obtained among the classrooms. The students in Ma's class apparently were more likely to emit behavior coded as physical contact/destruction. Such variation in the levels of behavior among classrooms underscores the importance of sampling peer behavior as well as that of a targeted student.

An additional point to be made regarding the individual data bears on the value of multiple behavior samples when assessments are made. If Ma had been observed only on day 2 of the observation sessions, one would have erroneously concluded that the subject differed from the peers only in terms of off task behavior. The subject's behavior on the three other categories was either equal to or less than that of the peers. However, that inference is contradicted in the data from multiple samples. Similar contradictions may be found in each of the other cases. The conclusion is clear. When attempting to document behavioral difference through classroom observations, multiple observations are essential if erroneous conclusions are to be avoided.

Utility

The usefulness of any classroom observation system similar to the one presented here is finally determined by attempting to apply it to the solution of individual problems. In this closing section of the paper,
an example is presented of how the system is used to establish the need for, and to evaluate the effects of, a social behavior intervention. Use of the observation system and development of the intervention were accomplished by a special education resource teacher as a part of that teacher's job responsibilities.

The problem. Marilyn was eight years old and in a second grade classroom. Her teacher, Mrs. Nelson, referred Marilyn to special services because "she was always out of her seat bothering other students and seldom completed her morning seat work during morning reading instruction." Marilyn's mother was contacted by the principal to determine whether she felt a problem existed; she told him: "I do not call you when a problem occurs at home, and I would appreciate it if you did not bother me at work with school problems." When Marilyn was interviewed to assess her view of the situation, she stated she did not like the school, the work, or Mrs. Nelson who she believed disliked her. Further, she claimed to have no friends in the room.

The assessment. Marilyn was observed in class for five days using the observation system previously described. Her behavior was observed to be 1.5 times greater in the "off task" category, and 1.7 times greater in the "out of place" category (see Figures 5 and 6).

The intervention. Once the resource teacher summarized the data he concluded that while Marilyn's behavior was apparently different from that of her peers; the data also showed that the entire class was emitting both categories of behavior at a high rate. The resource
teacher showed Mrs. Nelson the data and she responded by saying that this particular group of children had been especially hard to manage and that it had been a difficult year. She said, further, that when Marilyn had entered the class two months previously, she was just one more problem child. She finally concluded that "maybe the whole class needs some control."

Together, Mrs. Nelson and the resource teacher developed a simple plan to influence both Marilyn and her classmates. Briefly, the management system consisted simply of having the students earn the right to a play time by completing seat work assignments without leaving their seats or talking loudly. Each occurrence of either leaving the seat or loud talking resulted in a two-minute reduction in play time.

The results. To assess the effectiveness of the intervention, the resource teacher continued to observe the behavior of Marilyn and her peers. The results of the intervention on the behavior of both Marilyn and her peers are revealed in the data displayed graphically in Figures 5 and 6. As can be seen, the effects with respect to "out of place" behavior were marked for both Marilyn and her peers. The median frequencies for both during the second week amount to approximately one occurrence during the observation period. No such clear effect can be seen for "off task" behavior. A small reduction appeared to occur in Marilyn's off task behavior but no clear change occurred for the peers. At this point some revision might have been made in the program to influence off task behavior; however, the school year was ending and no further revisions were made.

The case study well illustrates the potential use of the observation system. Other teachers have used the system similarly, and some have made
more extensive use of the "other" behavior category to obtain data on behaviors not specifically defined. The case presented here was selected because it underscores the importance of assessing a child's behavior relative to the child's peers. Were the resource teacher to have assumed that Marilyn's behavior was the primary problem and have observed only Marilyn's behavior, the intervention subsequently developed with Mrs. Nelson might have been far different -- focusing exclusively on Marilyn rather than the entire class. The data obtained during initial assessment provided a clear basis in fact for the discussion between Mrs. Nelson and the resource teacher which led to a change in the teacher's behavior, which, hopefully, was fairer to Marilyn and benefited everyone involved.

Final Notes

The data obtained through using the system described here can be summarized and presented in a variety of ways. How the data are summarized is, of course, less important than how they are interpreted. Great care should be taken to use the data fairly for all persons involved. That means that once differences in behavior are empirically established, the importance of those differences in each individual case must be addressed by all parties with a vested interest in that difference (i.e., student, parents, teachers, etc.). To quote Robert Mager (1970) on this point, "If a discrepancy can be ignored it should be." A person who is out of his or her "place" more is not ipso facto behaving wrongly. A careful examination should be made at the same time of the extent to which the expectations for behavior are reasonable and appropriate. Within the perspective presented here, if we can eliminate the difference between the level of behavior which is desired by someone, and the
level of behavior actually emitted by the student, "the problem" has been eliminated. The most economical solution, then, is to set aside the desired or expected performance level and accept the performance actually emitted. Unfortunately, such "setting aside" is not usually the easiest solution for whomever holds that expectation for behavior.
References


Foster, G. C., Ysseldyke, J. E., & Reese, J. H. "I wouldn't have seen it if I hadn't believed it." Exceptional Children, 1975, 42, 469-473.


Gallup, G. "I'd really have seen it if I hadn't believed it." Exceptional Children, 1975, 42, 469-473.


Footnotes

1 Although the system described here is directed toward classroom social behavior, the procedures are generalizable and can be used anytime one wishes to obtain a measure of the discrepancy between two sets of behavior, whether those behaviors are academic or social, and whether they occur inside or outside of school settings.

2 This conclusion can be placed in perspective by referring to the peer rates of these behaviors presented earlier in this paper.
Table 1

Mean Frequencies of Behaviors for Target Students and Peer Samples, and Number of Target Students Exhibiting Two or More Times the Peer Frequency for Each Category

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Mean Frequency (per minute)</th>
<th>Number of Target Students Two or More Times Discrepant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target Students</td>
<td>Peer Samples</td>
</tr>
<tr>
<td>Out of Place</td>
<td>.15</td>
<td>.04</td>
</tr>
<tr>
<td>Off Task</td>
<td>.39</td>
<td>.11</td>
</tr>
<tr>
<td>Noise</td>
<td>.25</td>
<td>.17</td>
</tr>
<tr>
<td>Physical Contact</td>
<td>.03</td>
<td>.02</td>
</tr>
</tbody>
</table>
**Figure 1**

**RECORDING SHEET**

**Recorder:** [Name]  
**School:** [School]  
**Child(ren):** [Name(s)]  
**Date:** November 26, 1978  
**Circumstance:** Regular classroom

| Behavior      | Noise | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | Total |
|---------------|-------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|-----|
| **Noise**     |       |   |   |   |   |   |   |   |   |   | 10  | 10  | 10  | 10  | 10  | 10  | 10  | 10  | 10  | 10  | 10  | 10  | 40   |
| **Out of Place** | 0   | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 3    |
| **Physical Contact** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 2    |

**Peers**

| Behavior      | Noise | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | Total |
|---------------|-------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|-----|
| **Noise**     |       |   |   |   |   |   |   |   |   |   | 10  | 10  | 10  | 10  | 10  | 10  | 10  | 10  | 10  | 10  | 10  | 10  | 50   |
| **Out of Place** | 1   | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 3    |
| **Physical Contact** | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 3    |
| **Off-task**  | 0   | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0   | 1   | 0   | 1   | 0   | 1   | 0   | 1   | 0   | 0   | 0   | 10   |

25  2  7
Figure 2: Frequency of Pe and her peers on the four categories of behavior
Figure 3: Frequency of Ma and his peers on each of the four categories of behavior.
Figure 4: Frequency of Pa and his peers on each of the four categories of behavior.
Figure 5: Off task behavior for Marilyn and her peers
Figure 6: Out of place behavior for Marilyn and her peers
The Institute is not funded for the distribution of its publications. Publications may be obtained for $2 per document, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be prepaid.

Requests should be directed to:

Editor
Institute for Research on Learning Disabilities
223 Fraser Hall
106 Pleasant St. S.E.
University of Minnesota
Minneapolis, Minnesota 55455


*Ysseldyke, J. E., & Thurlow, M. L. Specific investigations to be completed during years two and three (Monograph No. 4). June, 1978.


*As part of its continuation proposal, the Institute was required to prepare these monographs. Because they are part of the proposal, they are not available for general distribution.


