Behaviorally anchored rating scales (BARS), a new quantitative method of employee performance evaluation, is advocated for teacher evaluation. Development of a BARS consists generally of five steps: a representative sample of potential raters generates the scales; the group identifies the broad qualities to be evaluated; the group formulates general statements representing various levels of performance for each quality and lists examples of behaviors at each level; a second group, representing the ratees, receives lists of the broad dimensions and of the specific behaviors and assigns the behaviors to the dimensions; raters then assign unused behaviors to describe satisfactory and unsatisfactory employees across several dimensions. A discrimination index is computed for each behavior and point values assigned to each behavior within the dimensions. Items above a criterion level of point variability are eliminated. Problems may arise with the raters' predictive acumen or possibly with the permanent assignment of values for specific behaviors. As an example of a BARS used in education, the scale developed by Marianne Price for evaluating special education teachers in the Montgomery County, Pennsylvania, Intermediate Unit, is appended. This scale measures professional growth, communications skills, rapport with students, planning, instructional skills, recordkeeping, behavior management, and interstaff relationships. (MH)
USE OF THE BEHAVIORALLY ANCHORED RATING SCALE IN EVALUATING TEACHER PERFORMANCE

Robert J. Beebe
The University of Mississippi

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Beebe

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented at the Sixth Annual Evaluation Network Conference - Memphis, TN, September 30-October 2, 1980
ABSTRACT

The paper addressed the concept, features, format, development, and relative merits of using the behaviorally anchored rating scale in evaluating teacher performance.

The discussion provides basic information on a technique which is at the cutting edge of industrial performance evaluation. Application of the behaviorally anchored rating scale technique to teacher performance evaluation is illustrated.

Suggestions as to how such a technique might be applied in a school district are offered, as well as critical comments intended to keep the technique in perspective.
USE OF THE BEHAVIORALLY ANCHORED RATING SCALE
IN EVALUATING TEACHER PERFORMANCE

As you are aware, numerous techniques have been
developed, and are in current use, for the evaluation of
employee job performance. There are various ranking
procedures, including straight ranking, alternative ranking,
paired comparison, and forced distribution techniques.
There are qualitative methods, including critical incident,
weighted checklist, and forced choice techniques. There are
management by objectives approaches. There are also direct
indexes, including measures of productivity, and measures
of withdrawal, such as absenteeism and turnover.

By far the most popular methods of employee per-
formance evaluation are those quantitative methods known as
rating methods. [For a fine discussion and comparison of
the various types of performance evaluation techniques,
see Cummings and Schwab, 1973.] In recent years a new
quantitative method of employee performance evaluation has
developed, and has attracted considerable attention in the
literature. This is the behaviorally anchored rating
scale, or BARS for short. In the years since 1963, when
the concept was first introduced, more than 125 journal
articles have appeared in which the authors discuss, develop,
test, or advocate the use of behaviorally anchored rating
scales in evaluating employee job performance. (See Schwab
and Heneman 1975; also DeCotiis, 1978).
Given the pressures both from within and without school districts for reliable and valid teacher performance appraisal mechanisms, it may surprise you to learn that only one application of the behaviorally anchored rating scale technique to public school teaching is reported in the professional literature. (Price, 1979) Both service and business positions have been analyzed, but not in the public schools.

Let me begin my discussion of behaviorally anchored rating scales with a bit of their history. In 1963, Patricia Cain Smith and L. M. Kendall took note of the extreme demands placed upon the quality of ratings in many situations. They stressed that such ratings should possess reliability both across raters and situations. Both the levels and components of such ratings should be clearly understood by those assigned the task of making ratings. It is only under such circumstances that raters can make valid ratings, perhaps more originally, be expected to use rating scales with conviction or agreement. The necessary concensus among raters can be achieved only if the raters themselves define, in their own terms, the kind of behavior which represents each level of each discriminably different characteristic, and which dimension of behavior is illustrated by each kind of behavior. The behaviorally anchored rating scale is Smith and Kendall's answer to this problem.
In a few moments I shall show you several behaviorally anchored rating scales developed by Marianne Price in her doctoral study for the position of Special Education Teacher. I want first, however, to describe the procedure under which such scales are developed. Consideration of these two elements separately was inspired by Dickinson and Zellinger, 1980. It is through this procedure that the chief advantages of the behaviorally anchored rating scale are achieved. It may be noted in passing that the procedure for the development of behaviorally anchored rating scales has been studied informatively independent of the behaviorally anchored rating scale format itself.

Smith and Kendall describe the procedure for the development of behaviorally anchored rating scales as similar to that employed to ensure the faithfulness of translations from one language to another. Material is translated into a foreign language, and then retranslated by an independent translator into the original language. Where "slippage" is found to have occurred, corrections are made. You will see how this retranslation simile applies to the present context in a moment. The behaviorally anchored rating scale development procedure consists generally of the following five steps:

1. A representative sample of the persons who will ultimately serve as raters are enlisted to conduct the basic work of generating the behaviorally anchored rating scales.
In education, it would seem that this reference group could logically include not only administrative raters, but also teachers. The point of calling together the work group is not to generate administrative solidarity, but rather common understanding and commitment. Both sides of potential evaluation problems ought, in my judgment, to be involved in scale development from the outset. We have a long history in education of such involvement.

2. The group identifies and lists the broad qualities or characteristics to be evaluated. The most strongly supported dimensions are selected for further analysis. Usually the number of dimensions centers about eight. The participants' own terminology is retained in identifying the dimensions. Let me show you what such a set of dimensions might consist of.

SHOW OVERHEAD TRANSPARENCY #1. (RJB note the source)

3. The group then formulates general statements representing definitions of high, acceptable, and low performance for each dimension, and generates examples of behaviors at each level for each dimension. These are edited into the form of expectations of specific behavior. Let me show you an example of a set of behaviors indicative of one dimension.

SHOW OVERHEAD TRANSPARENCY #2. (RJB block off left columns.)
Note: It may be worth noting here that not all investigators have identified and defined the broad qualities, characteristics, or dimensions first. Some have begun with specific behaviors, which are then grouped into appropriate categories.

4. A second group of persons, representative of the group sampled by the original work group, is provided a listing of the broad dimensions, and a second list of the specific behaviors developed by the first group. They are asked to assign the behaviors to the dimensions. Behavioral examples are eliminated if there is not a criterion level of agreement on the dimension to which behaviors ought to be assigned. Qualities or dimensions are eliminated if there is not a criterion level of agreement in the behaviors which ought to be included.

5. Other judges, perhaps the first group, are asked to use the remaining behaviors to describe the behavior of satisfactory and unsatisfactory employees as across the several dimensions. A discrimination index is computed for each behavior. Judges are also asked to assign certain point values to each behavior within dimensions, according to its desirability. Items above a criterion level of point variability are eliminated. Mean point values for surviving behaviors are retained and used in the behaviorally anchored rating scale format.
Let me turn now to the scale format. You will recall the dimensions of performance I showed you a moment ago. Let me show you several of the scales representing these dimensions. SHOW OVERHEAD TRANSPARENCIES #3, #4, and #5, if time permits.

The function of the scales is of course to assist in the evaluation of employee job performance. Let me describe how the rater is expected to perform this responsibility. Here I must confess to less guidance in the literature than has been present with respect to my prior comments. Nonetheless permit me to offer a few observations.

The rater is expected, over some no doubt specified period, to observe the employee's performance on the job. On each broad dimension of the appropriate set of rating scales, the rater assigns a point value to his or her rating of employee performance. The specific point values awarded are in accordance with the similarity between the behaviors observed and those for which points are specified on the behaviorally anchored rating scale. The specific comparison process followed, and this point is quite clear in Smith and Kendall's article, operates as follows. The rater forms a clear image of the employee's work, then predicts which of the behaviors listed in the scale the individual would be most likely to display.
It is here, in my judgment, that the behaviorally anchored rating is set adrift. The anchor drags in the mud of the rater's predictive powers. There is a second flaw, as I see it, which makes such anchors' value questionable. This is the permanent fixing of point values for the specific behaviors used as anchors. My experience both as ratee and rater suggests that few teaching behaviors maintain the same value from moment to moment, much less from teacher to teacher or class to class.

I recognize that the comments which I have just made sound contradictory. How can one complain both of the unreliability of point values in a scale, and then of the false precision in the scale? I suppose that the best defense is a good offense. I propose that we not sacrifice either for the other. A revision of the behaviorally anchored rating scale format, and corresponding minor changes in the scale development procedure would, I believe, preserve the best of both.

The specific behaviors generated to illustrate the dimensions agreed upon ought not to be sifted out on statistical criteria. The value of the behaviors is not in their statistical properties, but rather their representativeness of actual teaching behavior. The larger number of behaviors generated could serve as a guide to observation, and could provide behavioral anchors, but not for numerical ratings, but rather for the exercise of judgment on the part
of the rater. Perhaps some dichotomization of positive and negative behaviors could be performed to simplify the job of rating, but this step is beside the point which I am trying to make.

The behaviorally anchored rating scale is not a free lunch. Yet it may provide enough nourishment for us to continue the search productively for improved methods of teacher evaluation.
REFERENCES


BEHAVIORALLY ANCHORED RATING SCALES FOR EVALUATING
THE SPECIAL EDUCATION TEACHERS OF THE
MONTGOMERY COUNTY, PENNSYLVANIA,
INTERMEDIATE UNIT

1. PROFESSIONAL GROWTH
2. COMMUNICATION SKILLS
3. RAPPORT WITH STUDENTS
4. APPROPRIATE PLANNING
5. INSTRUCTIONAL SKILLS
6. RECORD KEEPING AND REPORTING
7. BEHAVIOR MANAGEMENT
8. INTERSTAFF RELATIONSHIPS

DEVELOPED BY MARIANNE PRICE
1978
1. PROFESSIONAL GROWTH

Seeks to improve instruction through the exploration of suggestions, new ideas, materials, approaches, etc.

After identifying her weaknesses, teacher takes additional course work to further her knowledge.

Teacher attempts new approaches to teaching and evaluates their success and failure.

Teacher has several behavior problems in class. She attends conference concerning positive discipline approaches to seek new behavior modification techniques.

A teacher consults professional journals in her field and implements some of the innovative ideas contained therein.

A teacher experiments with a new idea that was presented at an in-service training conference.

A teacher attends a conference in her field and when she returns she does not attempt to implement any new ideas/techniques with her students.

Even though she received in-service training on a new reading program, the teacher will not use it because it involves a change in her planning.

Teacher attends a convention but only attends evening social activities.

2. COMMUNICATION SKILLS

The ability to exchange information with administrators, professional persons, parents, and other adults

A classroom teacher asks a speech therapist for information regarding the nature of a particular child's speech problem and how she can help. The therapist talks with the teacher about the problem and prepares a list of activities that the classroom teacher can implement to help the child improve his speech.

Teacher keeps in close contact with counselors, or ancillary staff to keep all programs working toward the same goals.

The teacher holds regular conferences with parents, at their convenience. If parents are unable to come to the school, she makes a home visit.

Teacher consults with other professional staff in determining appropriate expectations for each child in the area of behavior.

The teacher confines her conversation with parents to her area of expertise, the classroom and instruction, and thus avoids treading into areas in which she lacks credentials and certification.

A child does not have his homework, so the teacher calls the mother and suggests that the child simply forgot to bring it to school, rather than accusing the child of not having done it at all.

Teacher does not inform supervisor of major changes in a child's educational plan.

The teacher told the parents all year long thatJohnny was making good progress. In the June conference, the teacher told Johnny's parents that she wanted to have him reevaluated because his progress had been unsatisfactory.

The teacher ignores a parent's request for a conference.

3. RAPPORT WITH STUDENTS

The ability to develop an accepting relationship with children, which facilitates their success and growth in a learning environment.

The teacher observes that a child who has difficulty expressing his feelings in front of others is extremely upset about something. She also knows that this child is embarrassed to let the other students know he is upset. The teacher summons his help in running off and assembling dittoes in the office. This provides an opportunity for the child to talk to her.

The teacher extends to children the same common courtesies that she expects her children to give to her, e.g., please, thank you, excuse me, etc.

A student criticizes the teacher's style of discipline. The teacher listens to his opinion and tells the student why he has chosen this method of discipline.

When working with distraught adolescents who are having a rough time communicating with and understanding parents, a teacher relates her past experiences and problems in the same area.

The teacher avoids excessive familiarity with students so that he/she is perceived as the teacher and not a peer.

A teacher unknowingly mispronounced the child's last name: The child soiled himself daily for six weeks until the teacher finally said his name correctly, at which point the soiling of his clothing ceased completely.

Upon completion of instructional session, teacher criticizes student for making numerous errors, rather than focusing on what he completed correctly.

During instructional session, teacher remarks, "You'll never be able to learn this," when the student makes an error.

4. APPROPRIATE PLANNING

The ability to organize instruction in order to achieve short and long-term objectives based on the needs and level of performance of each child.

A language therapist reviews all testing data on a child and outlines her goals for the child. As her program progresses, she constantly reevaluates her goals and objectives to make sure they are still appropriate to the child's needs.

The teacher is able to write an educational prescription appropriate to the needs of each child.

The teacher has her lesson plan prepared and all materials that she will be using in a given lesson assembled and at hand before the lesson begins.

Teacher determines a student's entry level in a reading program on the basis of diagnostic test information.

Teacher organizes books, pictures, art materials for discussion on national holidays which occur during the school year.

A teacher has identified the long-term goals for the child. She knows generally what she wants to do to achieve these so she does not develop any short-term objectives.

Teacher has not planned adequate amount of work to keep a student occupied and learning. As a result this student is bothering everyone else.

The teacher assigns the same dittoes as seatwork to all children in her class despite the fact that children's skills range from the first to the fourth-grade level.

Teacher establishes neither long-range goals nor short-term objectives for her students.

5. INSTRUCTIONAL SKILLS

The ability to implement the educational program through the use of appropriate materials and techniques.

The teacher has commercial materials that are too advanced for her class. She modifies them for use with individual students where possible. She also has numerous materials that she has made herself to fit individual needs.

In order to maintain the attention and interest of students, the teacher varies the pace of her instruction beginning with pencil-and-paper activities which demand much concentration on the part of the student and moving to a more relaxing activity such as an instructional game when she notices that the child's attention for the pencil-and-paper task is waning.

After a child is taught a particular skill, the teacher reinforces this skill.

The teacher is instructing a student in a second-grade level math book, but the student has not committed the basic addition and subtraction facts to memory. The teacher assigns student to an activity center where he practices these facts on a teaching machine.

The teacher uses the weekly magazine, News Pilot, to discuss current events and various social studies and science-related topics with her students.

The teacher follows text in sequential order, making daily assignments to each pupil. Help is given each pupil as it is needed. No supplementary materials are used. No instruction per se is given.

The teacher uses audio-visual materials, e.g., films, tapes, to fill up time, with no thought being given to their relationship to the instructional program.

While teaching beginning reading, a teacher uses teacher-made dittoes on which the printing is very small and unclear.

6. RECORD KEEPING AND REPORTING

The ability to keep accurate and current records, using this information for comprehensive reporting as required.

Reports are always objectively written and substantiated by records; they are submitted within time requirements.

The teacher documents (in writing) the instructional goals and objectives she holds for each child.

The teacher gives a detailed account of when a child has a tantrum, including the time of the tantrum and the behavior that preceded the tantrum.

The special education teacher documents the communications she has with regular class teachers about special education students.

When asked to justify Johnny's placement in a third-grade math text, the teacher is able to present Key Math Diagnostic Test results and the textbook placement tests to substantiate placement at the third-grade level.

The teacher keeps a daily attendance record, collects absence information, and hands in attendance report every two weeks.

After administering various testing instruments to a student, the teacher places the results in the student's file.

A teacher claims she has "tried everything" to get Johnny to do his assignments but "nothing works." The teacher's supervisor asks to see a record of what has been tried, for how long, etc., so that she can offer direction to the teacher. The teacher has no written record of what has been done and can only give the supervisor vague recollections regarding each type of intervention that was tried.

Teacher writes reports which include emotionally laden unsubstantiated comments.

The ability to establish and maintain a behavioral atmosphere that is conducive to learning.

The teacher established a behavior modification system involving group and individual goals for classroom behavior. Varying types and rates of reinforcement are established for individual students.

The teacher is alert to situations that can erupt into disruptive behavior. She takes steps to avert trouble before it occurs.

The teacher identifies student behaviors which interfere with the total learning process. She implements a behavior modification system utilizing reinforcers appropriate to each student.

The teacher uses a token economy with child who does not respond to social reinforcement.

A student has been evidencing disruptive behavior in the classroom in the form of calling out and interrupting. The teacher initiates a system for taking turns in speaking and asking questions.

Teacher rewards positive behaviors of children, ignores negative behavior.

The teacher takes away privileges to control classroom behavior.

A teacher yelled at a student in a classroom because he misbehaved. The child ran out of school.

Pupils frequently swear and engage in name calling. The teacher reprimands them each time this occurs. No decrease in the inappropriate verbal behavior is noted.

Verbal and physically aggressive behavior frequently occurs in the classroom. The teacher provides no consistent consequences to these behaviors.

8. INTERSTAFF RELATIONSHIPS

The ability to build and develop positive relationships with I. U. and/or district staff.

8.0

A hearing therapist volunteers her services to a school on an Open School Night for Parents program by offering to speak, test hearing or whatever services the principal or other teachers may want of her.

7.0

The teacher volunteered to "serve" an extra duty for a regular staff member who became ill.

6.0

When the principal gave permission to have a ditto machine placed in the speech therapy room, the therapist said, "I would like to do my job effectively. It is impossible to conduct speech therapy classes effectively when the ditto machine is being used. Let's work out a schedule for use of the ditto machine."

5.0

The I. U. special class teacher adheres to the time schedules and the policies of the school in which she works.

4.0

Special class teacher maintains good public relations with regular class teachers by eating lunch with them and conversing with them on a personal as well as professional level.

3.0

2.0

I. U. Self-contained teacher in district building does not follow normal channels of communication within the building.

1.0

An itinerant teacher fails to inform the school office that she is ill and will be absent.