DOCUMENT RESUME

ED 077 146 EA 005 258

AUTHOR	Beegle, Charles W., Ed.; Brandt, Richard M., Ed.
TITLE	Observational Methods in the Classroom.
INSTITUTION	Association for Supervision and Curriculum Development, Washington, D.C.
REPORT NO	ASCD-17948
PUB DATE	73
NOTE	95p.; Edited Papers presented at ASCD Southeastern Regional Conference
AVAILABLE FROM	Association for Supervision and Curriculum Development, 1201 Sixteenth Street, N.W., Washington, D.C. 20036 (Stock Number 17948, $3.50)
EDRS PRICE	MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS	Classroom Environment; *Classroom Observation Techniques; *Conference Reports; English Instruction; Instructional Improvement; Mathematics Instruction; Nonverbal Communication; *Public Schools; Research Methodology; Self Evaluation; Student Behavior; Student Participation; *Supervisor Qualifications; *Supervisors; Teacher Administrator Relationship; Teacher Behavior; Teacher Supervision; Teaching Styles

ABSTRACT

The ASCD Supervision Council established three regional conferences to (1) assure that supervisors in elementary and secondary education would have the opportunity for self-renewal through a wide variety of meaningful experiences, and (2) help supervisors and curriculum workers develop skills and competencies that could be shared with teachers. The Southeastern Conference had as its theme, "Systems for Observing In"School Operations." The papers presented at this conference are included, in edited form, in this booklet, and include Leslee J. Bishop, "Systems for Observing In-School Operations;" Robert S. Fleming, "The Supervisor as an Observer;" James Raths, "Problems Associated with Describing Activities;" Richard M. Brandt, "Toward a Taxonomy of Observational Information;" Donald M. Medley, "Measuring the Complex Classroom of Today;" Charles M. Galloway, "The Nonverbal Realities of Classroom Life;" Larry S. Bowen, "Use of the Flanders Interaction Analysis System;" I. V. Ahnell and Horace C. Hawn, "Self-Evaluation Through Video Tape Recordings;" William C. Lowry, "Content Analysis of Mathematics Instructional Materials;" Joseph E. Strzepek, "Analyzing the Content of English Instruction: A Point of View;" Harold R. Strang and James R. George, "Instrumentation in Monitoring and Recording Human Behavior;" and Richard M. Brandt and Hugh V. Perkins, Jr., "Observation in Supervisory Practice and School Research." (Author)

Observational methods in the classroom

Edited by
Charles W. Beegle and
Richard M. Brandt

Foreword by
Harold G. Shane

Association for Supervision and Curriculum Development
1201 Sixteenth Street, N.W., Washington, D.C. 20036

Price: $3.50
Stock Number: 17948

The materials printed herein are the expressions of the writers and not necessarily a statement of policy of the Association.

Library of Congress Catalog Card Number: 73-80535

• ACKNOWLEDGMENTS

FINAL EDITING of the manuscript and publication of this booklet were the responsibility of Robert R. Leeper, Associate Secretary and Editor, ASCD publications. Technical production was handled by Nancy Olson, Lana Pipes, and Teola T. Jones, with Mary Albert O'Neill as production manager.

Contents

Foreword *Harold G. Shane* iv

Introduction *Charles W. Beegle* v

1. Systems for Observing In-School Operations
   Leslee J. Bishop 1
2. The Supervisor as an Observer *Robert S. Fleming* 13
3. Problems Associated with Describing Activities
   James Raths 20
4. Toward a Taxonomy of Observational Information
   Richard M. Brandt 23
5. Measuring the Complex Classroom of Today
   Donald M. Medley 35
6. The Nonverbal Realities of Classroom Life
   Charles M. Galloway 45
7. Use of the Flanders Interaction Analysis System
   Larry S. Bowen 56
8. Self-Evaluation Through Video Tape Recordings
   I. V. Ahnell and Horace C. Hawn 61
9. Content Analysis of Mathematics Instructional Materials
   William C. Lowry 65
10. Analyzing the Content of English Instruction: A Point of View
    Joseph E. Strzepek 70
11. Instrumentation in Monitoring and Recording Human Behavior
    Harold R. Strang and James R. George, III 75
12. Observation in Supervisory Practice and School Research
    Richard M. Brandt and Hugh V. Perkins, Jr. 79

Contributors 85

• Foreword

FOR A NUMBER of years the field of supervision has been shaking free of traditions built up between the 1920's and the 1950's. One of the milestones denoting change was the late Kimball Wiles' *Supervision for Better Schools* (1950), a book which proposed ideas and concepts and which has stimulated thought and shaped practices for 20 years past.

Observational Methods in the Classroom extends the rim of our vision and updates our knowledge of contemporary trends in the field on which the 16 contributors focus their combined experience. As I reread the final manuscript copy, the material once again impressed me with its anticipation of tomorrow's supervisory practices. Here we find a nicely blended approach to sound, provocative ways of working with teachers and with one's associates in the timeless effort to improve instruction.

Dr. Beegle's useful and comprehensive introduction precludes the need for me to comment further on this important cluster of papers—presentations originally stimulated through the ASCD Supervision Council and the regional conferences arranged under its leadership. I am sure that readers will join me in thanking the authors for skillfully summarizing and extending the state of the art of supervision in U.S. schools.

Indiana University
May 1973

HAROLD G. SHANE, *President 1973-74*
Association for Supervision and Curriculum Development

Introduction

SUPERVISION, as it is currently known in typical school situations, seems to be perceived as of little practical use in improving instruction at the classroom level. Findings in various studies convey this implication, whether the perceptions being analyzed are those of teachers, administrators at the building level, or, indeed, supervisors themselves. Usually these findings are indicated in the form of a single, broad statement such as: "There is widespread role diffusion in the supervision of instruction." Yet generalizations at this level of abstraction are of relatively little help in charting new directions for: (a) designing more effective training programs for supervisors; (b) improving competencies of supervisors already in the schools; or (c) developing a more adequate theoretical base for the field of supervision.

Many school people hold that the goal of the supervisory leadership team of an American school is to help the teacher achieve a primary and authentic commitment to the intellectual, emotional, and physical development of children. The supervisor, therefore, must be a person who is self-renewing and inquiring, utilizing new findings for studying and knowing children, their lives, and their aspirations in the world.

In light of the evidence, the supervisor needs to find ways to help the teacher ask questions about newer or older methods of instruction. Will this learning experience produce what Maslow would classify as a growth-fostering environment for the child, or will it be growth-inhibiting? Does the learning environment encourage the child to become more open to stimuli about him, or does it cause the child to close, to build a psychological screen around himself? Does the experience help fulfill the needs of children? Are the realities of schooling congruent with stated beliefs and intentions so diligently conceived and printed in school philosophy and curricular programs?

Naturalistic observation could help in congruency testing. The importance of naturalistic observation lies not merely in providing the

informal grounds for daily decisions. It also assists the local practitioner in testing, in his own setting, the findings and procedures of basic research in order to see if, and to what extent, they are applicable.

The ASCD Supervision Council established three regional conferences to assure that supervisors in elementary and secondary education would have the opportunity for self-renewal through a wide variety of meaningful experiences. Conferences were established to help supervisors and curriculum workers develop skills and competencies that could be shared with teachers. This was done in the hope of learning the extent to which congruency is obtained between the intentions and the realities of schooling.

"Systems for Observing In-School Operations" was the topic for the Southeastern Conference. Papers and work sessions dealt with the areas of values, problem and design considerations, and observational methodology in order to develop skills needed for congruency testing of the intentions and realities of educational programs. The papers that were presented at this conference are included, in edited form, in this booklet.

The presenters, who viewed observational methodology from a values, problem, and design orientation, include the following:

• Leslee J. Bishop, "Systems for Observing In-School Operations." Concerns were expressed about current practices, sorties were made in and out of research, and some implications were drawn from developments presently under way. Bishop raised fundamental questions: What are the parameters of "in-school"? Do we have a perspective and instruments to assist, catalog, and analyze learning and growth? Are we capturing the essence of "in-school" experience?

• Robert S. Fleming, "The Supervisor as an Observer." Fleming presented five suggestions to help supervisory personnel use observational procedures more effectively in the future. Already partially tried and tested practices suggested include: (a) making a series of team observations, each with a separate, clear-cut purpose; (b) making shadow studies of individual pupils; (c) gathering a unique type of accountability data on selected pupils; (d) developing and working on simulation assignments, based on teachers' problems; and (e) making diagnostic observations. These suggestions can lead to improved opportunities for principals and supervisors to sharpen their own skills in observing educational practice.

• James Raths, "Problems Associated with Describing Activities." This presenter identified some of the problems requiring solutions in designing and carrying out observational studies. Among these problems

are the selection of important units of observation and the determination of appropriate sampling procedures. He suggested several possible components of a classification scheme for describing the worthwhileness of educational practices.

• Richard M. Brandt, "Toward a Taxonomy of Observational Information." Brandt presented the basic overview of various types of observational data used in the study of ongoing school processes. He discussed the strengths, weaknesses, and suitabilities of rating scales, checklists, narrative techniques, and various subtypes of each of these general procedures. He presented an overview of a key chapter appearing in his recent book on naturalistic research.

• Donald M. Medley, "Measuring the Complex Classroom of Today." Medley described two studies of changes in teaching style of teaching-interns. Changes were measured by applying OScAR to kinescopes of classes taught early and late in the first semester of teaching. Significant shifts in teaching style were apparent on two or three out of eight major factors assessed in each study. The utility of comprehensive coding systems like OScAR is clearly demonstrated for gathering solid observational data simultaneously on a number of key aspects of classroom instruction.

• Charles M. Galloway, "The Nonverbal Realities of Classroom Life." Galloway stated that some persons assume that nonverbal behavior represents the essence of relationship language and that the meanings communicated by it are often sharply different from verbal behavior. He reviewed a wide range of procedures for studying nonverbal behavior as a necessary adjunct to the study of verbal behavior. He also described several programs in which teachers have been successfully sensitized to the meanings which their own nonverbal behaviors carry for students and to detecting nonverbal information transmitted by students. Training programs for helping students understand the significance of nonverbal cues are also described.

• Larry S. Bowen, "Use of the Flanders Interaction Analysis System." Bowen pointed out the efficacy of using IA as a part of pre- and in-service teacher training. He had found this method to be quite beneficial in a summer institute. Using IA gives the teacher an opportunity to view his verbal behavior, which can, in turn, help him modify his interaction with students. Bowen stated that, once one has been trained in using IA, the process is relatively easy to use and the results are of value to both the teacher and the supervisor.

• I. V. Ahnell and Horace C. Hawn, "Self-Evaluation Through Video Tape Recordings." Ahnell and Hawn reported on some recent research that was conducted at the University of Georgia. Although the data gathered after one year of study revealed no conclusive evidence, they suggested ways by which a student might benefit from viewing himself on video tape.

• William C. Lowry, "Content Analysis of Mathematics Instructional Materials." Lowry demonstrated how "level of behavior" can and should be related to "content" in the selection of mathematics materials. Supervisors should be able to benefit from the suggestions included in this paper because the suggestions represent serious attempts to make material selection less subjective.

• Joseph E. Strzepek, "Analyzing the Content of English Instruction: A Point of View." In this presentation, Strzepek assumed the position that students, as well as teachers, should be trained in analyzing and evaluating the worth of an English course. He cited several well known educators who have dealt specifically with evaluating English curricula, in both the cognitive and affective domains. Because of the nature of communication skills, the suggestions made by this writer should be helpful to teachers of English. Moving beyond prescribed course content is essential for the total evaluation of an English program.

• Harold R. Strang and James R. George, III, "Instrumentation in Monitoring and Recording Human Behavior." Strang and George stated that, along with instructional materials and equipment, observational hardware is gradually being introduced in some institutions in recent years. They also reviewed the types of instrumentation now available for monitoring and recording human behavior. Many devices may not be appropriate for ordinary supervisory practices; however, in exemplifying a wide variety of observational methodology, certainly the device of instrumentation is essential.

• Richard M. Brandt and Hugh V. Perkins, Jr., "Observation in Supervisory Practice and School Research." Brandt and Perkins summarized some of the main themes emerging from the various papers which describe the state of the art in observational methodology. They recognized the value of such methodology for research purposes and cautioned against its possible misuses in supervisory practice. They highlighted its great potential as a means for improving instruction, learning, and self-fulfillment; and they advocated joint teacher-supervisor application of systematic observations in studying general problems of school life.

A variety of research was presented, from checklists to the utilization of observational hardware. All the methodology presented would not be applicable in a particular school setting; however, an attempt was made to survey a wide range of recent research in the field of observational methodology.

A supervisor is needed who makes a difference; one who acts effectively in maintaining and renewing teacher competence and the instructional program; one who can release the powers of teachers in the advancement of the instructional program. To make a difference, a supervisor cannot rely on position, but must gain power through giving evidence of expertise in his field. This is no easy task. If the instructional supervisor is to release human potential for the improvement of the educational system, his skills to help others look at congruency between stated intentions and classroom realities must be developed. These conference papers were prepared in order to enable supervisors to become more keenly aware of tools to help teachers systematically collect data from "in-school" operations.

CHARLES W. BEEGLE
Associate Professor of Education
University of Virginia, Charlottesville

Charles W. Beegle

Richard M. Brandt

CHARLES W. BEEGLE is presently Associate Professor of Education at the University of Virginia, Charlottesville. He has taught at both the elementary and secondary levels in public schools in Ohio, where he also served as a junior high school principal, junior high school consultant, coordinator of instruction (K-12), and Federal Project Director. His major professional interests include curriculum, supervision of instruction, community school education, and environmental education.

RICHARD M. BRANDT is Professor and Chairman of the Department of Foundations of Education at the University of Virginia, Charlottesville. His principal experience includes classroom teaching in public schools in Michigan, and college instruction at the Universities of Delaware, Maryland, and Virginia. He has written numerous articles and monographs related to child growth and development, pupil assessment, and observational methodology. Dr. Brandt's book *Studying Behavior in Natural Settings* (New York: Holt, Rinehart and Winston, Inc., 1972) has been a major contribution in the field of observational methodology.

1 • Systems for Observing In-School Operations

Leslee J. Bishop

WHAT is meant by the term, "in-school"? Reasonable, sensible, pragmatic persons know they cannot manage all the problems in all the arenas of life, much less solve them. So realistic parameters have to be set.

Yet "in-school" surely has become an anachronism. We know that most stimuli and subsequent motivations and learning are now as powerful out of school as in school. And the critics would say the out-of-school inputs are more relevant, more real, and more powerful (Reimer, 1972). Likewise, we know that sources of debilitation are other powerful forces coming from outside of school, whether they arise from cultural conflicts, genetic factors of intelligence, nutritional deficiencies, or life-style considerations.

Whatever their source, whatever their valence, these external and cultural forces exist within the school as stimulants or detractors. And we must contend with them. A home cannot contain a child, a school cannot educate a learner, a neighborhood cannot limit or control the inputs to a child's attitudes or knowledge. These are the products of the larger learning environment, of our political, social, and electronic ecology. Therefore, any approach to observation, learning, or analysis must be sufficiently flexible and adaptable to consider the entire environmental matrix.

So when we say "in-school" let us be realistic, and let us be predictive; because "in-school" is a changing concept. A learning society that is now considering four-day work weeks, four-quarter school terms, and the flow of ideas from satellites, computers, and cassettes will insist that "schooling" be a nonbuilding function; that the educator be concerned with education, not just schooling; and that we participate in the planning, transacting, analyzing, and evaluating of the educational gain that the learner achieves—wherever and however it takes place.

More and more basic information and skill items will be nonschool learnings. They will be the product of:

- Educational television (community, open circuit, and cassette)
- Environmental inputs from a better educated, more involved society
- Modular skill and information elements now being developed by commercial and consortional efforts
- Electronic and other facilities soon to be available to pupils at learning centers or homes.

In this situation, soon to be with us, school will be the place for diagnosis, evaluation, and socialization. Many of the transactions on which our instruments and classroom methods are now based will be nonschool-building events. Even school elements may well have to be observed, analyzed, and evaluated differently as school aides, differentiated staffing, and community resource sites and personnel become more functional.

Time, and Place, and School

You have only to recall the impact the planting season has had upon school schedules to anticipate the changes predictable with a shorter work week, mobile facilities and populations, and a concern for competence, not courses.

Thus our criteria, our processes for observing teachers, learners, and operations must include the near future as well as the immediate past. Where, how, and what a learner learns, how or where and by whom instruction is given, where and by whom learning is managed—all these require new and forward-looking procedures.

Program budgeting, population mobility, teacher teaming, learning packages, technology, economic restrictions and shorter time elements required by nongrading, four-quarter (or even more) segments of the year (Coleman, 1972): all these suggest that time, a major organizational element to date, will be restructured. This will force changes in our record keeping; we will view students in *specific* operations rather than more casually over time. So we will need more precise instruments and computers to record what is happening and what should happen next. Even so-called flexible scheduling will soon be as archaic as is our continued use of the basal text, yearly promotions, group evaluation, and courses organized and designed primarily to perpetuate bodies of knowledge—as though we did not have libraries, telephones, TV sets, and access to the riches of assembled and worldwide knowledge.

The task of education, and of educators, is to search out and utilize every learning situation, every learning site. We can, therefore, envisage school as a series of institutional and social agencies that range in responsibility from (a) mass education as now being promoted by the American Cancer Society, the National Heart Association, National Educational Television, and *Sesame Street*; (b) to a closer series of referral agencies and clinics that evaluate and prescribe, and that are concerned with health, safety, family planning, consumer protection, and economic survival; (c) then to another circle of influence such as education-related industries producing new materials, school-community learning centers, and education-related learning agencies that enfranchise educational experiences, that sell programmed learning, and speed reading, and inexpensive manuals that range from how to understand current economic policy to how to lose weight (Marien, 1972).

Perhaps all we are talking about in the whole learning matrix is that portion of educational experience which will someday be more like the intensive care unit in a hospital. Education is a process, not a place.

The economic crunch, the press for accountability, the plea for responsiveness, the forces of decentralization and power sharing, and the economic, social, and electronic alternatives to schooling will force a more open, more flexible, more time limited, and a much less child-care center concept upon us, as school *per se*. School will be different from a building concept, so "in school" will also be different.

In order to be safe, then, we talk about "in school." But a concern for parameters also requires that we reconsider our perspectives and our instruments, to decide whether we are developing procedures for assisting, cataloging, and analyzing learning and growth, *or* whether we are capturing the essence of in-school experience. These two concepts are not the same. Whether we are, for the sake of research and security, locking ourselves into concepts of a passing rather than a becoming institution, and whether we are continuing to be schoolmen instead of educators are not the same.

Place—and Media

A cartoon idea, expressed in many ways, has often created laughs where it should have caused concern. Frame one shows a teacher talking or lecturing to a class. Frame two shows the teacher replaced by a tape recorder or a TV. Frame three shows the tape-TV "teacher" and, in the seats where the students formerly sat, there are now tape recorders. No one is present, and there is evidence that the interactive, personal elements have now become mediated and impersonal.

This situation or its equivalent is occurring with increasing frequency. My point is not that we lament the occurrence, but that we analyze the situation. A number of implications are obvious. The one I want to stress is not that the classroom situation may have become sterile or inhumane, but that *the place and point of interaction has been changed.*

The roles of the teacher, the school, and the classroom have been changed. Interaction has moved from the classroom to whatever place or time the student chooses; it has moved from the classroom and the teacher to a message system where man-machine interface is an element in the transaction. Hopefully, such methods will clear the classroom of mere information transmission, and replace it with opportunities for intercommunication, for laboratory and personal interaction activities. The elicited response or behavior requires a new pattern or technique for observation, for gathering evidence regarding impact or effectiveness.

Observation and Technology

The technologists with their inputs and outputs boxes, the PPBS advocates, the accountability monitors, the systems analysts, the EDP flow charters, the management operators, and the like all have the momentum on their side. They represent a long overdue concern for analysis, for cause-effect calculations, and, more important, for outcome, especially learner gain. These developments require that we sharpen our objectives, select our means, and ascertain what was transacted and what was achieved with much greater precision and care than our generalized institutional approach has provided in the past (Provus, 1970).

These people and their techniques are not our enemies—they are colleagues, friends, and co-workers. As Brandt (pp. 23-34) points out, we need their tools and approaches to comprehend and manage the complexities with which we are confronted. We need their techniques to show how the affective is a gain as well as the cognitive, and as Galloway emphasizes (pp. 45-55), how the nonverbal can transmit signals as well as the verbal; and how the higher levels of thought, *à la* the taxonomies, need to be programmed into experience and achievement.

So the domains and the taxonomies likewise become our friends as they help us consider balance, as they provide elements to analyze and evaluate the range of critical learnings and growth elements. If we use them to fuse, and not fragment information and experience, they can move us forward.

New Chemical and Psychological Approaches Need Investigation

Public and professional interest has been great regarding some of the highly psychological approaches such as T-groups, encounter groups, and sensitivity groups. Likewise, the growing utilization of such techniques as behavior modification or behavior therapy suggests that we observe and participate in these efforts with both an objective and a critical stance. In a time of depersonalization, we have an obligation to seek out affective and effective ways of relating and reinforcing. In an age of criticism and lack of credibility, we still have the obligation to ensure individual integrity and privacy. The line between shaping and enabling is a thin one at times; the difference between information needed for individual relevance and concern for individual security may not always be clear. The professional roles in diagnosis and prescription require expertise, sensitivity, and communication of a high order.

Also, for example, the uses of drugs for stimulation, retention, and behavior control are both promising and dangerous. The primacy of the psychologist as the guru of behavior has to yield in part to the chemist. We can no longer afford to ignore the learning disabilities of nutrition and physical differences, even as one recent study suggests we are moving toward a genetically stratified society.

All the more reason, then, for developing and utilizing instruments that record our concerns, our constraints, our procedures, and our outcomes or gains. All the more reason for keeping ourselves and our publics informed with good data, with professional utilizations.

As society, and as an institution, we have reached a point of no return. For, once we start playing God with nature, with atoms, with social, genetic, and economic structures, once we start precisely organizing and programming ourselves, we have to go all the way by providing the elements or the "slippages" in the system that permit nonprogrammed elements to exist.

The nontechnological, the noncompetitive, and the withdrawn groups that are forming, along with the present passion for nostalgia: these are symptoms of a scared society which, like the adolescent, sees the overwhelming and new ahead so, for the moment, regresses and becomes the child again, seeking secure arms and tranquility—but knowing deep inside that in time this temporary hold must be rejected in order to achieve adulthood in development. So we, like the child, subconsciously try to "psych" ourselves for the indistinct future.

Thus in school, when we begin to be precise with all the pieces, then we must program what we want. Otherwise that which is easy to

program becomes our immediate future, for example, low-level cognition —easily developed, taught, and evaluated; or we emphasize only the basic skills because they can be monitored and tested. Yet we must also program for activity, for contemplation, for creativity (Bruch and Torrance, 1972). These too have to find a place in time and structure and instrumentation, or else we lose our professional souls again to parsimony and particulars.

> Men must be judged by the challenges they define for themselves. So far, they have attached more importance to the challenge of adventure than to the challenge of compassion, more importance to the challenge of technological grandeur than the challenge of human growth, more importance to the challenge of war than the challenge of peace, more importance to the challenge of productivity than the challenge of perspective, more importance to the challenge of the scientific intelligence than the human spirit (Cousins, 1971, p. 20).

Observation and Change

Educators tend to make the new curriculum, the new process, the new hardware an add-on feature; we add a course, a teacher, a piece of equipment, a time. But the real need is usually a reallocation, a restructuring, a new configuration drawn to accommodate or accomplish the new feature.

As we use new instruments, as we develop new data, as we provide different perspectives, one critical element is whether or not we can build the new into the central operating features of the old. In other words, we must be sure there is both compatibility and a support system in terms of process, personnel, priorities, and budget. Else what happens is that we have a new add-on that provides novelty and excitement; the once-exciting add-on is lopped off and disappears without a public or a conscious decision being made.

Developments in the area of behavioral objectives and learning packets are relevant here. Further, I would add a plea for complexity. No one approach, no one instrument, no one change will modify an institution as gigantic as "school." The literature of change is clear on this point (Brodbeck, 1963). Likewise, there is a tendency with the finite objective and the piecemeal packet to fragment our efforts and the learning experiences of children. My concern is that we keep these elements in perspective, that we search for both figure and field. This is true both in our programs, curricular and instructional, and in the observations and experiences of individual learners. Piecemeal objectives and gains, reinforced by methods, material, and data, are important; but so are individual gains in patterning and composite structure building.

Making a life, or a living, requires both the immediate gain and the distant goal, considered in relationship to the individual and the blooming, buzzing confusion we call the world.

Statistician or Humanist

A test for us as supervisors is to check whether we support the individual or the maintenance of the normal curve; whether we represent the maintenance of the organizational structures and schemes (regrouping, grading, classifying, testing, passing, failing, enforcing)—all of which are relatively unimportant to learning, as research has amply illustrated—or whether we support enabling procedures, learning opportunities, human interaction, success, individual pacing, individual structuring and restructuring, development of self-image and worth, and individual perceptions that include reality, coherence, and commitment.

We need, in addition to observational instruments, even more delicate mechanisms or, as described by Simon (1967), "mirrors for behavior." A well-known TV comic has popularized the phrase, "what you see is what you get." So we get what we build in; we find what we set out to observe; we monitor the objectives we have programmed. Or, as Thoreau suggested, "But lo, men have become the tools of tools."

Yet the reinforcement studies; the nonverbal, the body language studies; the self-fulfilling, "Pygmalion" research tell us that what we transmit in expectation, what we reinforce by our looks and our attitudes, also teaches. ASCD has called some of this the "unstudied curriculum" (Overly, 1970).

So even with complicated instrumentation we cannot dispense with the sensitive generalist, the humane observer who seeks to find awareness, propensity, motivation, learning styles, and creative properties that do not show up so directly on our scales, our instruments, or our validated checksheets. This is not a plea for mysticism or another psychological "black box," but a concern for empathy, for wholeness, for that which sometimes is interstitial among our calculations and program thrusts. Maybe I am also proposing that we continue to be aware that the heart really goes "lub dub," rather than the way it shows up on the electronic dial as "bleep bleep." The supervisor, like Marcus Welby, has to understand both.

If we assume that schools are for learnings, that curriculum is at best a plan, that media and technology are extensions and means of man, then we can ask,

What do we want our children to become? If we translate this question into somewhat more operational questions, these would include: What do we

want our children to come to value? What do we want them to be able to feel and see and hear and smell and touch? From what do we want them to learn to get pleasure? What do we want them to understand about themselves and the world of nature and man? How do we want them to behave toward other human beings? Toward what do we want them to be inclined to commit themselves? What technical abilities do we wish to cultivate in them? (Tumin, 1967, p. 78).

Revolution in Decision Making?

There is one major, almost revolutionary development that we want to watch. Possibilities for this revolution exist because of mediated learning opportunities, because of libraries and access facilities such as ERIC, because of the concern for immediacy and competency as contrasted to credentials and courses. Evidence can be found in the "open" universities, in the freedom schools, in the literature of the voucher plan, in the objectives of the performance contract, in the continuing struggle for quality in educational television and adult education, in concerns for decentralization. This revolution is obscured by court cases, busing, school consortia, tax struggles, and ad valorem taxes. This list of obfuscations could be indefinitely extended—and that is one of the symptoms of the problem.

The immediate future may well see not the state, not the local school board or school administration, not even the teacher, but rather the individual learner become the chief decision maker; the individual and his parents may well become the assemblers, the architects, the decision makers. To some extent this is already true, and always has been. But with individualization, media and access, and learning packages now as viable possibilities, this development is the most startling—and perhaps the most overdue—of all. (Even Skinner and Orwell are not excluded from this consideration.) So the situation can well become not "the teacher and his supportive staff," but "the learner and his supportive staff." Institutionally we will not let it happen, but the existence of this possibility will change the institution and the nature of decisions about the learner. For as we move from institutional certification to performance criteria (which do not require institutional treatment), we speed this movement.

As we open educational process and make it less a "black box" operation; as we include media, aides, and community; and as we view the operation with a concern for objectivity and not imprimatur as the standard, we hasten the advent of this possibility.

The Importance of Feedback

It is critical that the data we collect be reliable. That is an established concern—and in some camps it seems to be the major concern, and perhaps it is to the researcher. Validity is another question, if we consider validity in terms of educational purposes and processes, as well as particular instrumentation. This is one point where we need to pause and consider. It must be our determination that we do not reinforce that which ought not to have occurred in the first place.

The research and observation instruments have provided us with powerful tools, and with impressive data which we have repeated often but have not reconsidered adequately. For example, classroom interaction studies reported by Flanders (1970) have shown a high percentage of teacher talk and controlling behavior, have shown serious conflicts between verbal and nonverbal behavior. We have lamented these facts as evidence of a stress on talking and controlling rather than on pupil learning and participation. In the same vein, we were appalled at the quality of early TV, primarily because it consisted in large part of an emphasis on the "talking face." Yet we registered our concern as dissatisfaction with TV, rather than forging a determination to change the classroom approach and processes we saw mirrored so clearly in the TV program.

We need to build in the process of feedback; otherwise what we do is collect further evidence of inadequacy. Feedback is not just the data; feedback must be the data and the process for constructive reshaping. Feedback is not to preserve the status quo, but to define more adequately what we need to do. It cannot wait for a convenient time some year hence; it must be built in as the process now. Trial runs we need, field tests we need, but these are still inadequate because they tend to be one-time adjustments, often too late to change the materials, the format, or the instructional strategies. Thus, as we observe, as we collect, as we analyze, basic considerations still exist: how do we communicate the data, and to whom; and how do we modify the process, and by whom?

For example, what have we done with the data that:

- The low anxious child learns complex tasks more rapidly than the high anxious child
- But the high anxious child learns simple tasks more rapidly than the low anxious child
- Or for maximum—or appropriate—learning, the teacher ought to modify his or her teaching style depending on the goals of the particular lesson?

For the students who are roughly in the top third of their classes consistently, over a number of years, the school serves to build a positive self-concept and provides them with some immunization against later emotional stress. For these students, the school environment has a powerful positive effect not only in cognitive learning but also in the affective domain. However, for the students who are roughly in the bottom third of their classes, consistently over a number of years, the school serves to infect them with greatly lowered self-esteem and a lowered resistance to emotional stress (Bloom, 1971, p. 44).

We should be concerned with designing different methods of instruction as a way of handling individual differences, rather than concentrating only on differences in what the person is expected to learn or differences in pacing (Cronbach, 1971, p. 49).

Subjects which are required, sequential, closed, and which emphasize convergent thinking should, insofar as possible, employ mastery learning strategies (Bloom, 1971, p. 33).

There is nothing sacred about the normal curve. It is distribution most appropriate to chance and random activity. Education is a purposeful activity and we seek to have the students learn what we have to teach. If we are effective in our instruction the distribution of achievement should be very different from the normal curve (Bloom, 1971, p. 20).

I am distressed by the fact that we have accounted for pennies, but have not accounted for learning increments; that we join AAA for its road maps but reject a plan for direction and routing with regard to school change. I would like to see our curricular structure as clearly developed and visualized as the architect's vision of a house must become before it can be implemented.

We owe it to the students to make education as real as possible, to make it as relevant and as exciting as possible, to make it as precise and directional as possible, and to make it as individualized and as humane as possible. Maybe these things are not all feasible, but it is our obligation to strive for such goals, using whatever talents, processes, and resources we can marshal.

What Is Our Profile? Where Are We on a Continuum?

In the enterprise of education we inevitably take value positions. We describe in operations who we are, what education is, and what we want our pupils to become. One could draw up a series of continuums —or a series of concept confrontations.

On one such continuum, regarding *image*, some would see us engaged in a Don Quixote quest, hopelessly romantic, astride a mount

swayed by the weight of time and rapidly going to pieces as the aging elements grudgingly give up their ancient grip (Fabun, 1970). At the other end of this continuum we are the proprietors of the "little black box," enigmatic, calm, inscrutable, a process and a product that has already crept out of the future.

Or perhaps as we view our *curricular and instructional efforts* we see a different continuum; at one end is a prescription containing:

1. Finite objectives, drawn from taxonomies and domains
2. Validated media and materials
3. Shaping methods, experiences, strategies
4. Reliable observations and records
5. Precise measurement and feedback mechanisms.

At the other end of this line is a recipe, not a prescription. It proposes with Rousseau's *Emile* that we "keep your child's mind idle as long as you can," and the recipe can be found in a Hawaiian cookbook, entitled *A Cook's Tour of Kauai.* It is called simply "Preserved Children," and goes like this:

> Take 1 large field, half a dozen children, two or three small dogs, a pinch of brook, and some pebbles. Mix the children and dogs well together; put them on the field; stirring constantly. Pour the brook over the pebbles, sprinkle the field with flowers, spread over all a deep blue sky, and bake in the sun. When brown, set away to cool in a bath tub.

Education is a high priority, high stake enterprise. Without care, opening sensitive arenas for investigation is as likely to produce a Pandora's box as a panacea. And depending upon the analogy, it may be a challenge or a "can of worms." In any case, what we are engaged in here involves a look at where we are going and how we achieve our goals.

References

Paul M. Allen, William D. Barnes, J. L. Reece, and E. Wayne Roberson. *Teacher Self-Appraisal: A Way of Looking Over Your Own Shoulder.* Worthington, Ohio: Charles A. Jones Publishing Co., 1970.

E. L. Bentley and Edith Miller. *Systematic Observation.* Atlanta: Supplementary Educational Center, 1969.

Bruce J. Biddle. "Methods and Concepts in Classroom Research." *Review of Educational Research* 37: 337-57; 1967.

Benjamin S. Bloom. "Mastery Learning and Its Implications for Curriculum Development." In: Elliot W. Eisner, editor. *Confronting Curriculum Reform.* Boston: Little, Brown and Company, 1971. Copyright © 1971 by Little, Brown and Company.

Joseph C. Bondi, Jr. "The Effects of Interaction Analysis Feedback on the

Verbal Behavior of Student Teachers." *Educational Leadership* 26 (8): 794-99; May 1969.

Richard M. Brandt and Hugh V. Perkins, Jr. "Research Evaluating a Child Study Program." *Monographs of the Society for Research in Child Development* Vol. 21, No. 62; 1956.

May Brodbeck. "Logic and Scientific Method in Research on Teaching." In: Nathaniel L. Gage, editor. *Handbook of Research on Teaching*. Chicago: Rand McNally & Company, 1963. pp. 44-93.

Catherine B. Bruch and E. Paul Torrance. "Reaching the Creatively Gifted." *The National Elementary Principal* 51 (5): 69-75; February 1972.

John P. Campbell and Marvin D. Dunnette. "Effectiveness of T-Group Experiences in Managerial Training and Development." *Psychological Bulletin* 70: 73-104; August 1968.

James S. Coleman. "The Children Have Outgrown the Schools." *Psychology Today* 5 (9): 72-75, 82; February 1972.

Norman Cousins. , "Lunar Meditations." *Saturday Review* 54 (33): 20; August 14, 1971.

Lee J. Cronbach. "Comments on Mastery Learning and Its Implications for Curriculum Development." In: Elliot W. Eisner, editor. *Confronting Curriculum Reform*. Boston: Little, Brown and Company, 1971. Copyright © 1971 by Little, Brown and Company.

Elliot W. Eisner, editor. *Confronting Curriculum Reform*. Boston: Little, Brown and Company, 1971.

Don Fabun. *The Dynamics of Change*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.

Ned A. Flanders. *Analyzing Teaching Behavior*. Reading, Massachusetts: Addison-Wesley Publishing Company, 1970.

Michael D. Marien. *Alternative Futures for Learning*. Syracuse, New York: Educational Policy Research Center, 1972.

Norman V. Overly, editor. *The Unstudied Curriculum: Its Impact on Children*. Washington, D.C.: Association for Supervision and Curriculum Development, 1970.

Malcolm Provus. "Toward a State System of Evaluation." *Journal of Research and Development in Education* 3 (4): 87-96; Summer 1970.

Everett Reimer. *School Is Dead*. New York: Doubleday & Company, Inc., 1972.

Anita Simon and E. Gil Boyer, editors. *Mirrors for Behavior: An Anthology of Classroom Observation Instruments*. Vol. 5. Philadelphia: Research for Better Schools, Inc., 1967.

Sivasailam Thiagarajan. *The Programming Process: A Practical Guide*. Worthington, Ohio: Charles A. Jones Publishing Co., 1971.

Melvin Tumin. "Teaching in America." *Saturday Review* 50 (42): 78; October 21, 1967.

John Withall. "The Development of a Climate Index." *Journal of Educational Research* 45: 93-99; October 1951.

2 • The Supervisor as an Observer

Robert S. Fleming

RELEVANT TO the field of supervision are various observational techniques which I have developed and used over the years. These techniques make the role of the supervisor more viable and lead to increasingly meaningful curriculum activities. Observational procedures are realistic in that they lead to change of events or change of other activities; provide useful beginnings and useful data as feedback of a current system; and provide benchmarks around which future changes can be compared. I have five specific techniques which I shall propose, each with a clear-cut, single purpose.

Team Observation for Multi-Purposes

Suppose we were to arrange for a team of people to visit a given school. This group might include several principals, some supervisors, and a guidance worker. In some cases, teams have been composed of teachers. Let us assume that the team would take a week to do this task, and on each of five days we would have a group looking for a designated period of time for a specific item. For example, let us assume that the area of health, physical and mental, is an important area of concern.

Let us assume that we have a group of people who are free for a given block of time, say an hour, between 10 and 11 a.m. on a given day. Then we ask this team of people to take a walk to every spot in that school where the children are working. The members of the team should pace themselves so that in the indicated period of time they could cover the school. The questions they would raise might relate to the area of health and comfort. Are the chairs comfortable for the children? Is it too hot or too cold? Are the children dressed appropriately? Is anyone ridiculed or met with sarcasm? Is anyone rejected or excluded? Is anyone embarrassed?

There are, of course, all kinds of hazards; one hour is not sufficient

for the group to understand the situation. The approach may be varied or modified in any way that seems appropriate. Let us assume that in a given period of time this team has visited every spot in the school and has made independent observations.

At the end of the given period of time, the group members have returned to a central point and have been asked to summarize what they have seen. After they have summarized their observations, the group members might have six, seven, eight, or ten independent sets of observations of that school during that period of time. They then might pool this information and begin to come out with a series of observations that were validated by the fact that several had seen the same thing.

Now let us go on to the next day. Suppose on the second day between 10 and 11 a.m. the team takes a walk to look for materials. Questions might be raised: "What materials are being used in this school? How extensive are they? What are they like? What does it add up to? What do you see?" Following the observation, the group members repeat the same procedure, going back and summarizing, pulling out the common elements that all people have seen, and here they have a set of observations about materials.

The procedure is repeated on the third day using an item such as student involvement—we define it and describe it. The team goes out into every spot in the school, looking to see how the children are being involved in various activities, and the technique is repeated.

The fourth day the team might look for what we call "big ideas in the school today," big ideas in the curriculum that the children seem to be working on. The same procedure is repeated. On the fifth day the observation team might also look at how self-evaluation occurs: "How are children looking at themselves? How are they looking at the products of their work? What opportunities are there for them to participate in a discussion of how they are doing? How are they getting along? What is happening?"

During a period of five days, a team of competent people has taken a look at five different aspects of the school, and the team members have made independent statements. At the close of the five days of observation they develop a picture in terms of these five items of the school. A summary is then presented to the faculty. What kinds of priorities can we begin to establish? Here are some handles we might begin to hold. Observation is a basis for getting some handles in order to begin to look at ourselves.

An even better plan might have been for a principal and a team of his teachers to do the same thing. We did this once in a New Jersey school; and at the end of the week we plastered a room with our observa-

tions. No one had to evaluate the teachers because they had looked at themselves and their colleagues, and their observations gave them a basis for saying, "Let's come to grips with how we are getting along and begin to establish some priorities of what we do next." Observation becomes a beginning point.

Shadow Study

A second procedure might be used which I call a shadow study procedure. In a shadow study, we look at the youngsters themselves. For example, each observer could identify one tenth grade student early in the morning and go with this student wherever he goes all day long. He observes the tenth grade boy or girl in English, history, mathematics, French, physical education, home economics, etc.

The observer keeps two kinds of records, on a very simple page with a line down the middle, and with the date indicated. As he is observing this student during the day, we ask him at 15-minute intervals to describe on one side of the page what the setting is like, and on the other side of the page how did the student respond?

Suppose you had collected shadow study data for a number of students on a given day in a given school, then you had added up your summaries of what the situation was like and how these persons responded.

I imagine—and we found this in one of the studies—that the data would tend to show that, in that series of schools on a given day, the teacher talked all day long. The teacher talked in English, the teacher talked in history, the teacher talked in mathematics, and the teacher talked in science; and there were relatively few opportunities for students to get in on the act.

We also found that many of these youngsters had no opportunity for participation, for clarification, for raising questions, for probing, for experimenting, for crying out, for doing anything. They sat and they listened all day.

This is a simple kind of observational procedure in which we go directly to the students and we shadow them, and we get a body of information concerning these students in this school for the day. How do we use such data for sizing up the situation? We could pull all these bits of data together to find out what the school is like and what opportunities are provided. Again, this procedure is only a beginning, but it establishes a kind of data collection that could help in more effective educational decision making.

Modified Child Study Techniques

A third procedure relates to some work that is being done in Virginia. We are hearing a great deal about accountability. Some of the things I have been doing in Virginia for the past year relate to accountability in a variety of ways.

We asked each teacher to write three simple case studies or case summaries. First, we asked the teachers to select by their own criteria a child who could be described as making reasonable progress in school. Second, we asked them to select a child who was not doing well by their own standards; and third, to select one they would describe as a behavior problem. The teacher was asked to look at the accumulated record for each of these three children, to talk with him, to examine the products of his work, to talk with his parents, to do anything to provide for a simple, relatively short summary of each of these three children.

Suppose you had this information, on three children from each grade within a given school. The information compiled might have been a combination of observation and analysis of other available data concerning the child at the moment. Then some visitor—the principal, the supervisor, a guidance person, or some other person—is asked to spot some of these. The visitor comes in and visits for a while, then returns several times and observes. Take a look at this student. Read the summary. Do you see the same kinds of evidence? This is another observational procedure that might be important, in that you are not checking on the teacher, but are helping the teachers to refine their observational techniques, extending their ability to collect and analyze data. You are helping teachers to come to grips with the realities of in-school operations.

I have a feeling that the accountability story really has to do with how well each child is getting along, how well he is known, what information can be assembled about him. This is good observation procedure. Perhaps the foregoing discussion is a way of getting at another aspect of accountability.

Problem-Solving Observation

A fourth procedure is quite different. Suppose we had interest in some creative activities for a given faculty group. I got this idea from the book *Assessment of Men*,[1] which reports the procedures used during World War II to select some of the OSS candidates. It was found that

[1] Office of Strategic Services. *Assessment of Men.* New York: Holt, Rinehart and Winston, Inc., 1948.

certain qualities could not be measured with a pencil and paper test. Groups of men and women were placed in a field situation with a problem, and they then had to work out their own solution to the problem.

A problem situation could be established: 25 teachers, along with chairs and tables, are placed in a room which has in it several copies of today's newspaper—*The New York Times, The Washington Post,* and *The Richmond Times Dispatch.* The agenda for the group of teachers has to do with looking at today's newspaper. We divide the faculty into small teams, and each team is given an assignment. The assignment for the first team could be people, places, and big ideas in the news today. A second team might be given the want ads and the idea "What's for sale?" The third group might be given the sports section, the theater section, and the section dealing with the arts; then we might say, "What's to do?" Another group might be given just the sports section and be asked, "What's the game?" The idea is that the teachers take their assignment and think of many other instructional uses of the paper. They might resort to color, a skit, a dance. You see, I am getting at very different kinds of procedures.

As these people are at work, members of a team are observing them. They are observing the resources: What kinds of new resources are being utilized in this group? What kinds of competencies can be observed in this faculty group? What is the quality of imagination and insight they are bringing to bear in the solution of this simple little problem? Who is generating the idea and how could it spill over into the curriculum?

Here I am really searching, but I am trying to find some other ways of "turning on" a group of people. I have a hunch that we have too often in the past put people in a circle with an ashtray and a cup of coffee and said to them, "What are your problems?" Instead of asking them to identify problems, we should give them a problem, and the problem might be, "What's the game?" Maybe out of this can come some new kinds of impressions of the staff. You might identify some competencies in this faculty group which could be encouraged and could be turned on to more creative and more imaginative ways in which to work in the future.

Diagnostic Observation

My fifth plan is one which I call diagnostic observation. Many supervisors are skilled in a given area. When a teacher trusts the supervisor, when there is real acceptance and trust, the teacher might request

a visit to seek advice. He or she might say, "I am falling on my face with my job, I seem not to get through. How can you help?" So the supervisor comes in and observes very carefully. At the end of this session they sit down to think back through the classroom experiences. The supervisor may be able to raise some penetrating questions which in turn help the teacher put his or her finger on concerns, such as, "At what point did the lesson fall down?" "At what point was I in too big a hurry?" "At what point could other materials have been useful?" "Could I have worked differently with an individual and/or a group?"

A video tape might be helpful in this kind of observation. We might sit back and analyze the teaching-learning situation in terms of some of the preceding questions. Such a system is used to diagnose the situation, perhaps to validate some hunches, perhaps to provide data for future analysis. The tape might reveal further evidences of inadequacies, and yet also provide a starting point to do something.

I am still with the notion of a supervisor being the engager, and I still believe that this has to be done in a manner in which confidence and trust are specific to points of mutual concern. We have looked at a number of techniques of observation but the question remains: observations for what?

Another Look at the Five Techniques

Let us reexamine these techniques very briefly. The first, team observation, dealt with identifying handles to get a hold on problems, establishing priority, sizing up a school, and providing a comprehensive look at purpose at a given point and time. I think these are fairly reasonable and appropriate uses of the observations. The shadow study technique is useful in determining the impact of a day's program on a student: What is it like to sit there all day long? It also aids in providing for problem identification and/or hypothesizing about what is at stake in the school situation for this particular individual. It helps to check assumptions about the work and opportunities and the dynamics of a given school. The follow-up case summaries dealing with accountability provide a helpful way of beginning an evaluation of the school, increasingly comprehensive in the way in which we look at the purposes of the school, or a comprehensive evaluation of the work of an individual. They are also useful in looking at the validity of our school records and in spotting some of the tremendous gaps. I am of the opinion that this is one of the most urgent needs of most of our schools.

The fourth example that I used dealt with working our way out of a problem. This is helpful in identifying the teaching competencies of

the staff, in finding increased resources and creative potential among the faculty, and in determining more realistic expectations of what teachers can do. The last technique, diagnostic observation, helps in determining needs. This adds up to the fact that observations can serve as a means of focusing more attention on the totality of the curriculum of the school.

Observations can provide means of obtaining more information concerning the relevant factors of what we are now doing in a given school. Observations can be useful in supplying more data concerning processes that are being employed in a given setting. They are useful in giving more freedom and encouragement to individuals to innovate and change what they are doing. They can be a means of giving more emphasis to the learning which is occurring with the children. Out of all of these techniques could come a rethinking of the humanities in the school, and the replanning of specific areas, such as the social studies, the language arts, the mathematics program in a given school. Out of observation could come more attention to creative approaches to group life and to individual life.

And now, finally, how do we develop some of these observational competencies? Where do we go from here? This is a task which might well be assumed by teams from our institutions and our schools. Perhaps another focus for in-service education might be in this area. We might have different preparation programs for administrators and supervisors.

There are authors these days who write about *Death at an Early Age, 36 Children, The Me Nobody Knows, Crisis in the Classroom, Don't Smile Until Christmas,* and *Behind Classroom Goals,* and on and on. What about us? What about you and me? What do we do to vitalize our schools? How do we begin? How do we write a different version of the dynamics of something important happening in a given classroom? Perhaps there is a way out, and this way out is to look in our own backyard, to look in our own schools, to increase our own skills and observation, to identify points at which we might begin.

I think that we have a kind of disease these days, and it bothers me terribly. The disease is one of stressing the negative aspects of schools and teachers. Few people are actually rolling up their sleeves to get in and look, to get in and look with precision, and to try to identify those points at which change could be brought about. Much of what I am talking about does not require a government grant; neither does it require the action of the curriculum committee in our institution. I am talking about helping people to look and see, to listen and hear, to feel and to act.

3 • Problems Associated with Describing Activities

James Raths

ALL INSTRUMENTS and associated methodologies for collecting descriptive data share common problems. The sections which follow are intended to facilitate the identification of particular problems which face many of us who wish to collect data focusing upon classroom activities.

Defining the Unit. As with all classification systems, the problems associated with identifying the unit of observation seem common. Flanders used a unit of time (three seconds) and generally classified every three-second interval into one of 10 categories. What is a unit appropriate for classifying activities? If a teacher spends two minutes asking children about the latest moon shot almost prefatory to a lesson on the division of fractions, is that two-minute exchange an activity? During a free reading period, should the fact that a student read *Time* magazine for two minutes and *Sport* magazine for ten be recorded as two different activities? How can a log differentiate between a long-term assignment on which a student might work off and on for three weeks and a short-term assignment that takes most children approximately 20 minutes to finish?

Classifying the Unit. Once the unit of observation has been identified, a classification system needs to be developed that captures the characteristics of each activity that are deemed important. A category such as "requires the student to use a pen" hardly seems worth recording; however, the notation that the assignment places the student in a novel setting might indeed be provocative. The categories that make up the classification system must reflect the values that discriminate between, say, a free learning environment and more structured and authoritarian situations, if that is what is to be described.

The following characteristics, originally included in a larger collection published elsewhere (Raths, 1971), are advanced for consideration as possible components of a classification scheme for describing the

worthwhileness of educational practices. Needless to say, categories would need to be developed to classify all those activities which fail to meet the components.

1. The activity permits children to make informed choices in carrying out the activity and to reflect on the consequences of their choices.

2. The activity assigns to students active roles in learning situations rather than passive ones.

3. The activity asks students to engage in inquiry into ideas, applications of intellectual processes, or current problems, either personal or social.

4. The activity involves children with reality.

5. The activity asks students to examine in a new setting an idea, an application of an intellectual process, or a current problem which has previously been studied.

6. The activity gives students a chance to share the planning, the carrying out of a plan, or the results of an activity with others.

Selecting a Sample. A third difficult problem associated with monitoring the activities assigned in a classroom is that of sampling. The drawing of reliable samples, of course, is a difficult problem regardless of the observational system that is being employed. If a sample of assignments is drawn on Monday, is the resulting description characteristic of the ones given out during the remainder of the week? If the assignments distributed to a social studies class are classified, are the distributions of characteristics of those assignments similar to others given in mathematics, language arts, or science? Data need to be collected which indicate how varied are the assignment characteristics under study as settings, times, seasons, and subject matter context are changed.

Instrumentation. Ways of recording the assignments also need to be developed. Teacher logs might be an approach—but perceptions of teachers may not be totally accurate. The use of observer teams might be the most efficient, but impractical in terms of cost. Steele, House, and Kerins (1971) have reported high reliabilities using student descriptions of activities going on in their classrooms. Whatever the source, whether it be teacher self-report, recordings of observers, or student descriptions, efforts must be made to assess the reliability of the measures emanating from the procedure.

Data Reduction. Once the mass of information is recorded, ways must be found to summarize the data in some meaningful form. For instance, measures of central tendency or ratios of one sort of activity to another might be used. For communication purposes, it is crucial that

some rational approach be taken to encapsulate the observations taken on a set of classroom activities.

Concept Development. Underlying most of the questions raised in this section is the need to develop a new vocabulary to describe accurately those aspects of a lesson that do indeed reflect our beliefs. Speaking analogously, we need to develop concepts which aid our understanding curriculum to the degree that the notion of "germ" aided medicine in coping with contagion or that the term "atomic number" helped physicists in their comprehension of atomic behavior. Huebner (1966) called for such a development in deploring our sole use in schools of the "efficiency" model based on ideas of instrumentality that is so prevalent in current discussions about schools and schooling.

References

Dwayne Huebner. "Curriculum Language and Classroom Meanings." *Language and Meaning.* Washington, D.C.: Association for Supervision and Curriculum Development, 1966.

James Raths. "Teaching Without Specific Objectives." *Educational Leadership* 28 (7): 714-20; April 1971.

J. M. Steele, Ernest R. House, and Thomas Kerins. "An Instrument for Assessing Instructional Climate Through Low-Inference Student Judgments." *American Educational Research Journal* 8 (3): 447-66; May 1971.

4 • Toward a Taxonomy of Observational Information

Richard M. Brandt

THIS PAPER will present the beginnings of a taxonomy of observational data, that is, the types of information an observer can gather in order to conduct systematic, empirical investigations of ongoing educational processes. A more extensive review of observational data gathering procedures is presented elsewhere (Brandt, 1972a).

Ratings

By far the most widely used form of behavioral data is the rating. A rating represents an estimate of the degree to which a particular characteristic is manifest. Thus, a youngster may be checked by his teacher at report card time as being above average in *responsibility* and *independence*, average in *cooperativeness*, and below average in *industriousness*. A rating scale is generally presumed to represent a continuum from complete absence to complete presence of a given trait.

General Ratings. Rating scales take many forms. Some scales are merely straight lines with descriptive terms placed along them at regular intervals (none, some, much, all) to help the rater place his X's in the spots which correspond to his judgments. Some lines have only two or three dividing points; others have five, six, seven, or even more. Rating scales seem disarmingly simple to construct and fill out, requiring only the naming of traits to be rated and then the mere placing of a check mark to represent one's estimates of how much these traits are in evidence.

As accurate measures of the persons or traits being assessed, however, many ratings are virtually worthless. For scientific purposes their validity is often highly questionable and, for making decisions about people, they can be not only misleading but highly damaging. Frequently they reflect more about the subjective state of the rater than the true

nature of the ratee. Subjective biases are particularly enhanced, furthermore, when characteristics to be rated are rather global in nature and ambiguous in meaning; traits such as *dependability* and *self-reliance*, for example. Such traits may carry quite different meanings from one person to the next.

Rater biases are also accentuated when ratings are presumed to cover a considerable period of time and a wide variety of conditions and ratee behaviors. The common practices of rating pupil characteristics at the end of a six-week marking period or of supervisors rating teachers only two or three times a year can produce no real information about the person being rated except the impression he has made on others. Such impressions not only lack precision in describing the persons or qualities in question, but they are frequently dead wrong. Even conscientious teachers can remember only a handful of specific examples of a particular youngster's behavior during that time when they are reaching their decisions regarding what categories to check on his report card.

Some teachers are more prone than others, furthermore, to spread their ratings of their pupils over the full range of the rating scale or to weigh them in a positive or negative direction in terms of how socially acceptable is the characteristic in question. In addition, when several traits are being rated at once, a general positive or negative attitude toward particular pupils, called the halo effect, tends to minimize trait differences so that a child whom a teacher likes is customarily rated high on all virtues and low on vices regardless of the specific nature of the quality in question. This lack of discrimination from quality to quality increases the overall positive feedback some children receive from the school and the negative feedback others receive, driving a wedge between the "good" and the "bad" kids.

Illustrative of this powerful tendency not to discriminate between qualities being rated is a study by Cook and Richards (1972). Factor analyzing principal and supervisor ratings of 236 teachers on 23 characteristics, they found that only two factors accounted for 61 percent of the variance, namely one representing the supervisory biases and the other, those of the principals. Two superhalo effects rather than real differences in teacher behavior were the primary findings. Instead of using 23 separate rating scales, almost the same information could have been obtained "by simply instructing an evaluator to rate the teacher as unsatisfactory, satisfactory, good, or excellent."

These disparaging remarks are not to suggest discontinuance of our use of ratings. Many important human characteristics need to be evaluated, and ratings represent the best method available. Other areas than those that can be measured by tests are in need of assessment, and

1. This child typically gives in to others, fails to assert his own opinions, avoids trying new things, is reluctant to participate in group activities, refuses to enter competitive situations. He gives up when criticized. He seeks constant reassurance.

2. This child gives in to others, lets others speak for him, follows rather than leads, stays on the sidelines, enters activities cautiously, and does not seek competitive situations. He becomes discouraged and/or defensive when criticized. He frequently seeks reassurance.

3. This child participates fully and confidently only in those activities with which he is familiar and in which he has experienced prior success. He usually becomes defensive when criticized. He often seeks reassurance.

4. This child participates fully in most activities—expressing confidence in his abilities, enjoying opportunities to try new things. He sometimes becomes defensive when criticized. He responds positively to reassurance when in challenging situations.

5. This child participates fully in most activities—expressing confidence in his abilities, enjoying opportunities to try new things. He frequently laughs at his own mistakes, and he accepts criticism as a challenge. He seldom requires reassurance.

1	2	3	4	5
Billie Jones Martha Smith	Sue Babcock Tom Martin Bill Olson Marjorie Pine	Dick Adams Bob Cook Phyllis Doyle Bob Nelson Ronald Rolland Mary Thompson Sally Waters	Phil Baker Mary Carter Mark Elder Jim Foster Sue Gallagher Joe Hendricks Jack Little Grace Piper Alice Reynolds Bob Swanson Don Thompson	Bill Davis Merle Johnson Sam Mann Mavis Olds Tom Ronson Norm Zigler

Figure 1. Teacher Rating of Her Class on a Behavioral Description Scale (Brandt, 1972a; p. 123)

ratings can be a potentially valid means if certain precautions are followed. I shall suggest a few:

1. Trait ambiguity can be reduced by providing illustrative operational descriptions which distinguish between the specific kinds of behavior included and excluded from the meaning of particular characteristics.

2. All members of a group should be rated on one characteristic before the next trait is considered. When the second trait is taken up, ratings of the first trait should be disregarded. This procedure will minimize halo tendencies. Because this is somewhat more cumbersome than the common procedure of rating many qualities of the same person at one time, it may necessitate limiting those qualities to be assessed to a few important and behaviorally distinct ones.

Figure 1 illustrates how a class of primary age children could be rated on a single scale in which the points along the trait continuum are defined behaviorally. Notice that no descriptive term, such as self-confidence, is used to identify the scale, in order to minimize raters' tendencies to respond to their personal interpretation of such a term rather than to the behavioral descriptions indicated.

3. Ratings of student writings, art work, and other products should be done blindly, that is, with names covered up and other identifying marks removed. This can be accomplished with almost no extra effort and should be done by teachers routinely in grading pupil work.

4. Situations can be preselected in which particular traits are likely to be demonstrated, and ratings can be made at the time of observation rather than days or weeks later. Several separate ratings of the same type of situation, where a given trait is manifested, can later be combined, just as several test scores are often combined at the end of a marking period, in order to obtain an average rating over a larger time interval. Not only are the original ratings likely to be more valid than when done only at the end of a long time period, but the consistency and variability of behavior will be more evident.

In addition to the general types of scales already discussed, two other types of ratings should be mentioned briefly, each of which has commendable features.

Rankings. Ranking behaviors, people, or products according to particular criteria is a highly useful rating procedure. The football coach does this when he names his number one, two, and three quarterbacks. Ranking forces discriminations to be made among all items in a group (people or things) and, as such, makes it less easy for judges to make socially acceptable responses in each instance.

Nominations. Many human attributes are distributed in normal curve fashion, with only a few individuals standing out at each end of the distribution. Such exceptional qualities are easily noticed, and there

is likely to be considerable agreement among observers as to who possesses them.

One can obtain behavioral ratings of high validity, therefore, by having raters name people or items from a group that most exemplify a particular trait. For each of three traits (leadership, withdrawal tendencies, and aggressiveness) Havighurst *et al.* (1962) used several such items to elicit nomination data from River City pupils. For example, pupils were asked to write the names of classmates who best fit such descriptions as (p. 176):

1. . . . boys and girls who make good plans.
3. . . . boys and girls who stay out of games . . . don't like to play hard.
4. . . . ones who break rules—rules of the school and rules of games.

School life provides numerous situations in which classes select individuals to perform particular roles or functions. When the selection process is done by secret ballot, nominative data are obtained regarding the collective judgment of peers toward each other. Combining nominative data of this sort, collected on different occasions for different purposes, can provide highly useful records regarding the reputation and standing children have within their own groups.

Narrative Data

A second major type of observational information is that which I shall refer to generally as narrative data. In collecting this type, observations are made of ongoing events, and concurrent records are developed for later analysis and interpretation. The attempt is generally made to preserve behavior in essentially the same form and sequence in which it occurs, so that it can be dissected closely at a later date. Types of narrative data include (a) stenographic, anecdotal descriptions; (b) specimen records; (c) audio and video tape recordings; (d) ecological descriptions; (e) systematically obtained photographic records; and (f) anthropological, sociological, or psychoanalytic field notes. With the possible exception of the latter subtype, behavior is recorded in sequence as it happens and without interpretation. The essence of narrative reporting is the objectivity and completeness with which behavioral and setting information is recorded, so that resulting descriptions accurately represent that which actually takes place.

Observing and reporting such information does not occur automatically. People, including both teachers and supervisors, are inclined to see what they want to see and expect to see. Being able to report an

event in rather complete, objective fashion is a skill which comes only as one works hard to accomplish it. I have found that even well-trained psychologists have difficulty separating fact from opinion; that is, for example, describing an interview with a parent or a classroom observation of a child in objective, complete, and accurate language without inserting a good deal of their own subjective reactions as well.

The importance of objective, narrative recording is the permanent, relatively unbiased report that remains after an observation is over. With such a record, especially when it is combined with other records, analysis and reanalysis can be accomplished in systematic, scientific fashion. With subjective reports, on the other hand, it is never possible to go back to raw data for reexamination or reanalysis.

Not only is objective, complete observation and recording important for research endeavors but for ordinary professional activity as well. Teachers need to be able to notice and later describe children's behavior accurately in talking with parents and other teachers, in assessing what steps need to be taken to improve instructional programming, and even in maintaining understanding and supporting relationships with the youngsters themselves.

Similarly, supervisors would do well to learn how to make objective, faithful, and interpretatively neutral records of class activities they observe as a basis for later feedback and discussion with teachers. The classroom is a highly active place with well over a hundred teacher-child interactions occurring in a typical hour of class time. Much can go unnoticed, and much of what a teacher does occurs spontaneously, without opportunity for preplanning. The advantage of freezing some of this activity and interaction is the opportunity to reinspect it in a leisurely way, apart from the moment of decision, with the hope that new insights may be achieved in the process.

Not only can anecdotal descriptions be useful in this regard, for both in-service improvement and research, but modern electronic gadgets are now readily available for recording all kinds of instructional activity. Ahnell and Hahn indicate in their paper (pp. 61-64) how video tape recording (VTR) can lead to improved self-evaluation. Galloway (pp. 45-55) likewise makes use of VTR equipment in assessing nonverbal behavior. Medley (pp. 35-44) reports on the use of video kinescopes both as a training procedure and for the assessment of teacher change.

More simple and inexpensive to use, audio taping can produce highly useful records of classroom discussion patterns, pupil oral expressiveness, and teaching style. An hour's tape recording of a "show and tell" period turned out to be the most valuable single item out of three weeks of full-time observation I made recently of life in a British infant

school (Brandt, 1972b). This tape recording permitted the type of in-depth scrutiny needed to identify thoroughly the manner in which the teacher questioned children and drew so much good explanation and discussion from them. It was necessary to replay this tape several times before a meaningful set of categories could be constructed which would permit all of her statements to be coded efficiently and with high intercoder agreement.

Checklists

A third major variety of objective data is obtained with what I generally refer to as checklists. Checklists consist of category descriptions for behavior, events, or conditions that can be used to tally or otherwise record symbols standing for the specific types of behaviors or conditions observed.

Checklists can be used not only in direct observations, but also in the processing of narrative records. Narrative records are not particularly useful until they are processed in some systematic manner, either by rating whatever behaviors are recorded or by classifying them according to some appropriate category system.

Some time ago, Medley and Mitzel (1963) pointed out that, except for straight narrative reporting, an observer can make only two kinds of judgments in recording what he sees. One is qualitative judgment, referring merely to the existence or lack of existence of a particular attribute, with no attempt being made to estimate the degree to which it prevails. With this type, specific behavior is evaluated categorically as representative of a particular kind of behavior of which it is but one instance. The other type of judgment is an estimate of the degree or extent to which some specific attribute is manifested in the observed situation. The judgmental process is a rating, which has already been discussed.

While there may occasionally be some overlap in these two types of judgment, the first type (that is, categorical judgment) is usually used when a single situation is observed and classified. Ratings, on the other hand, are most frequently made in relation to a number of single behaviors or situations. Frequency or time counts of categorical records, of course, ultimately allow the second kind of judgment to be made also, namely the extent to which certain attributes are manifested over a number of situations.

Checklists take many forms. Several will be described briefly, with differences pointed out and possible applications suggested. Just as ratings have been overused, or at least misused quite often, observational

checklist data have not yet received the attention they deserve from school personnel.

The past decade or two of educational research has produced a wide variety of checklists that can be highly valid data-gathering instruments for teachers and supervisors to use in their daily activities. In addition, checklists can be constructed locally to meet particular needs one might have, once their potential utility is recognized and the general procedures understood for their development.

Static Descriptors. The term static descriptors refers to those relatively stable attributes of persons or settings that are to be checked systematically with each observation. They often serve as the primary basis for grouping observational data at a later date for analysis and interpretation. Sex, occupation, marital status, IQ, social class, and educational level are typical of many static descriptor variables that might appear on an observational form and would serve to identify the nature of the persons whose behavior is recorded.

Similarly, many situational features may need to be recorded routinely, such as the location of incidents, the number of persons present, time of day, type of content, or weather conditions. In school, such setting conditions as subject matter, type of ongoing class activity, and grouping arrangements often need to be checked, depending on the kinds of behavioral comparisons that will be of paramount interest. The main reason for their being listed and checked in each instance is so behavioral data will have a basis for comparison and contrast in later analysis.

Action Checklists. In recent years hundreds of action checklists have been developed for recording behavior as it occurs in precoded form. With a list of behavior categories in front of him, the observer tallies or records in symbolic fashion each behavior in which he is interested. For an action checklist to be useful, behavior must be readily classifiable and the categories making up the checklist must usually be mutually exclusive; that is, a particular behavior can be coded under only one category.

In the Flanders system, for example, whatever is occurring at a particular instant will be symbolized by only one number, the one which indicates the type of teacher talk or student talk (or possibly silence or confusion) which is then taking place. Galloway's category system permits types of nonverbal behavior to be coded also. Medley's PROSE system (Medley *et al.*, 1971) includes a variety of static descriptors and even some rating scales, and ongoing pupil and teacher behavior is recorded primarily with several interrelated action checklists, each with mutually exclusive categories.

Useful action checklists can be much simpler than those already mentioned. A class seating chart, for example, can be used to tally children's contributions during a social studies discussion, if one is interested in knowing only how many children take part and how often. A content analysis of teaching materials can be conducted also through simple classification and tallying, as Child, Potter, and Levine (1946) did with children's readers to see how many times storybook boys and girls do what they are told by adults, with positive or negative consequences following, and how many times they follow their own directions and with what types of consequences.

One rather simple action checklist was developed and used by Minuchin (1969) and her Bank Street colleagues to compare the types and amounts of pupil interaction of children attending traditional schools with those attending progressive schools. Children were observed one at a time and each interaction was coded as it occurred along four dimensions: (a) who initiated the interaction (teacher, child); (b) toward whom were child-initiated interactions directed (teacher, child, or other); (c) the manifest affect (friendly, neutral, or hostile); and (d) the purpose or orientation of the interaction (task-oriented, mixed, personal and/or social). Several days' use of this checklist showed children from the progressive schools to have proportionately more child-initiated contacts with their teachers, significantly fewer hostile exchanges with their teachers, and more frequent task-oriented, teacher-initiated contacts than those in traditional schools.

In addition to those systems requiring whatever behavior occurs at a given moment to be classified into one or another category, another type of action checklist is becoming popular, that which Medley and Mitzel refer to as a *sign* system. With this kind of system, all behavior is watched over a given interval of time (for example, 10 seconds, 5 minutes, or even 30 minutes). At the end of the interval, all selected types of behavior which occurred one or more times are checked once. This system permits watching for several distinct types of behavior during the same time interval.

Figure 2 shows a partial set of such behavior signs from the Medley *et al.* (1971) PROSE instrument. The observation time interval in this case is 100 seconds. During this interval, three of the listed signs were noted. Such a system is especially useful for recording events which do not occur frequently, but which are considered important enough to note whenever they do occur.

Much of what we do is categorical. If we attend a teachers meeting, for example, we cannot be going to class or home studying at the same time. Life offers choices, and at any particular moment what we are

	Used numbers
X	Used words
	Sang or talked to self
	Helped another pupil
	Asked another pupil for help
X	Had an accident and was hurt
	Showed fear
X	Cried
	Lost temper
	Told on another child
	Ridiculed by an adult

Figure 2. Partial List of Signs Recorded for One Child Over a 100-Second Observation Interval (Adapted from Medley *et al.*, 1971)

doing represents a resolution of various pressures to do one thing rather than another from the options available. A child will be task involved, engaged in social interaction, or daydreaming, for example, but not all three things at the same time. Action checklists are designed to keep a record of such categorical behaviors.

How well he does what he is doing, on the other hand, is a matter of either rating how task involved he is or keeping a fine enough breakdown of all the specific categorical behaviors he exhibits in the course of being task involved to permit such a judgment to be made later.

Activity Logs. Working days for teachers, pupils, supervisors, and even curriculum directors are divided into various periods by the major shifts in activity which take place. From 8:00 to 8:25 a.m., for example, a supervisor may be occupied primarily with reading his mail from the day before and dictating letters in return. From 8:25 to 8:40 several phone calls might be made to schools, establishing arrangements for a forthcoming in-service meeting with teachers. On through the day there are natural breaking points as one activity replaces another. One's scheduled plans and calendar entries do not always specify all of these periods, but the appointments instead. It is possible and often quite illuminating to keep an ongoing record of all activities in order to determine how time is actually spent.

It is surprising also how different the data derived from class activity logs are from those indicated in lesson plans. Teachers in the Baltimore Early Admissions Program kept logs for several weeks and initiated numerous changes in their teaching patterns as they themselves

came to realize how much time was being devoted to certain matters and how little to others. (See Brandt, 1972a, p. 106, for a copy of the log form.)

Discrete Event Records. A teacher or supervisor may wish to keep records of certain types of special events whenever they happen, such as contacts with parents, by listing the date of each contact, checking the nature of the contact (homework trouble, child behavior advice, teacher- versus parent-initiated, etc.), and perhaps making some brief narrative notes of what was said. The types of mistakes made on paper (spelling, punctuation, word usage, arithmetical error, etc.) can also prove useful when kept in the form of discrete event records. The key feature of discrete event records is a clear definition of the events to be recorded so they can be recognized the instant they begin, and complete and faithful notes can be kept of each one.

Standardized Situation Responses. Many school or classroom situations are essentially the same for all youngsters, even in a highly personal, well individualized program. Teacher assignments and class discussions obviously present youngsters with a basically similar stimulus complex. Records of pupil responses to such well specified, standardized situations can reveal a great deal about pupil individuality in attitudes, likes and dislikes, abilities, and even self-concepts.

Constructing such checklists is merely a matter of describing ahead of time the specific nature of an upcoming situation and listing the behavior options possible so they can be tallied during observations of children in those situations. Natchez (1959) found, for example, by checking a rather lengthy list of behaviors exhibited when children took their regular turn in their reading group, that youngsters with the lowest reading abilities also manifested significantly greater anxiety and frustration, as indicated by such behaviors as "looks at teacher when she comes to a hard word."

Contrived Situation Responses. When a teacher structures a class situation primarily to measure performance, a record is obtained of the pupils' contrived situation responses. It is quite possible to structure many observational situations in such a way that they appear to the responder as ordinary teacher requests rather than as tests.

Performance Records. Many tasks, games, and assignments are well defined, administered in a similar manner from person to person, and able to be scored in precise, objective fashion. Performance records kept on high school football teams, and players on those teams, represent the basic observational data used by opposing coaches to plan their strategies for upcoming games.

When teachers keep careful records of school assignments completed and the specific errors and accomplishments achieved on these assignments, they can go about planning their instructional strategies in similar fashion and, hopefully, with better success since the game they play is really more important. The essential ingredients of performance records are (a) a clear statement of each task and (b) a symbolic notation that indicates each success or failure of performers.

References

R. M. Brandt. *Studying Behavior in Natural Settings.* New York: Holt, Rinehart and Winston, Inc., 1972a.

R. M. Brandt. "An Observational Investigation of Instruction and Behavior in an Informal British Infant School." Paper presented at the AERA convention, 1972b.

I. L. Child, E. H. Potter, and E. M. Levine. "Children's Textbooks and Personality Development: An Exploration in the Social Psychology of Education." *Psychological Monographs* 60 (3): 1-54; 1946.

M. A. Cook and H. C. Richards. "Dimensions of Principal and Supervisor Ratings of Teacher Behavior." *Journal of Experimental Education* 41: 11-14; Winter 1972.

R. J. Havighurst, P. H. Bowman, G. P. Liddle, C. V. Matthews, and J. V. Pierce. *Growing Up in River City.* New York: John Wiley & Sons, Inc., 1962.

D. M. Medley and H. E. Mitzel. "Measuring Classroom Behavior by Systematic Observation." In: N. Gage, editor. *Handbook of Research on Teaching.* Chicago: Rand McNally & Company, 1963.

D. M. Medley, T. Quirk, C. G. Schluck, and N. P. Ames. *The Personnel Record of School Experience: A Manual for PROSE Records.* Princeton, New Jersey: Educational Testing Service, 1971.

P. Minuchin, B. Biber, E. Shapiro, and H. Zimiles. *The Psychological Impact of School Experiences.* New York: Basic Books, Inc., 1969.

G. Natchez. *Personality Patterns and Oral Reading.* New York: New York University Press, 1959.

5 • Measuring the Complex Classroom of Today

Donald M. Medley

ONE OF THE most exciting developments in educational research in recent years has been the appearance of many new instruments for measuring classroom behavior—instruments which make it possible to record observations of such behavior in objective, quantifiable form. The effect of the increasing availability of these devices and their widespread use has been virtually to revolutionize educational research. Yet I am convinced that their major impact will be felt through their use in the supervision of classroom teachers. I have elsewhere set forth my reasons for believing that this is true (Medley, 1971) and will not repeat them here; I will only remark that this impact will be felt in two ways:

1. The use of objective observational records as feedback devices is one way. An observational record is objective and therefore more easily accepted by the teacher than either the opinion of or a rating by a supervisor. It is usually more precise and detailed as well. And, finally, it tends to be fairer, or more comprehensive, than a subjective judgment. The supervisor who does not use an observation schedule tends to focus his attention on that part of a teacher's behavior which seems most in need of attention; one who does use such a schedule must attend to all aspects of the behavior. This puts the teacher's behavior in a better perspective and guards against neglect of areas in which problems have not yet come to a head but may soon do so.

2. The primary focus of this paper will be on a second important use that systematic observational procedures may have in supervision—and that is to measure changes in teacher behavior. It is generally accepted that pupils learn more rapidly when they can see results—when they are kept informed of their own progress. Is there any reason to doubt that the same thing should be true of teachers who are trying to increase their own skill? Assurance from a supervisor that one is showing improvement may provide some impetus to teachers to improve further.

But how much better would it be to have precise, objective measurements of changes in behavior from one supervised lesson to another than to have to accept someone's judgment—or should I say guess?

In the research in teacher behavior in which I have been involved in recent years, we have twice had the opportunity to assess the amount and kind of changes that take place in a teacher's behavior over a period of time. Because such studies are uncommon, I propose to describe some of the findings of these two—mainly to illustrate the potential that observational techniques have for detecting changes in teacher behavior over time.

A Study in Eight Dimensions

The subjects of the first experiment I shall describe were 54 undergraduate student teachers, all of whom did their student teaching in one campus elementary school in a large Eastern city (reported in Schueler, Gold, and Mitzel, 1962). The classrooms in the school were wired so that as many as three TV cameras could be mounted on the walls of any one of them and controlled as to pan, tilt, and zoom from a remote console. What the cameras saw could be recorded on a kinescope film continuously for up to 25 minutes. During each student teacher's first two attempts to teach a class, which occurred during the first two of the 14 weeks, one 25-minute kinescope was made of what went on; and during each student's last two attempts to teach the same class, which occurred during the last two of the 14 weeks, two more 25-minute kinescopes were made.

Variations in the behavior of the same teacher from one to another of the kinescopes made during the same two-week period were regarded as reflecting instability in teacher behaviors from day to day, subject to subject, etc. Differences between behaviors recorded 10 weeks apart, however, insofar as they exceeded short-term variations, were regarded as reflecting stable changes in teacher behavior.

After all the films had been made, trained coders viewed all four films of each teacher (in a random order and without knowing which film had been made at which point in time), and coded the behaviors on an instrument known as OScAR 3 (Observation Schedule and Record Number 3).

OScAR 3 was an omnibus instrument which was so long and complex that three coders were needed. One coder paid attention only to the verbal interaction, and coded each utterance into one of a set of 17 categories (which were basically an expansion of the ones developed

by Withall, 1949). A second coder watched to see how many of about 100 specific behaviors, or "signs," occurred during each three-minute interval. These signs were related to teaching techniques and emphasized cognitive processes. The third coder had a list of some 40 other signs to watch out for, mainly related to affective-emotional climate. In addition, at the end of each three-minute interval he rated the teacher on such things as voice, mannerisms, gestures, grace, and posture.

To reduce the number of variables, we factor analyzed the data derived from these codings and obtained eight orthogonal factors which preserved about two-thirds of the variance in them (Medley, 1964); these factors are named and described in Figure 1.

It is worth emphasizing that, since these eight dimensions were orthogonal, they were uncorrelated. In other words, a teacher high on one—say Presence—was no more likely to be high on another—say Warmth—than one low on Presence. Even if we arbitrarily decide to recognize only two levels on any one of them (high and low), this means that we can describe 2^8 or 256 distinct kinds of teachers using only these eight dimensions! (If we also distinguished "medium" teach-

NON-AFFECTIVE CLIMATE

A. Teacher Role

(1) *Presence.* Teacher keeps good order, is rated high on use of voice, gestures, etc.; uses clarification and neutral rejection.

(2) *Informative.* States objectives; relates lesson to pupil needs, past learning; high on information-giving statements.

(3) *Imaginative.* Provides for individual differences; makes use of apt, creative examples, techniques; arouses high pupil interest.

B. Pupil Role

(4) *Activity.* High questioning behavior on part of pupils and teacher both.

(5) *Initiative.* Variety of pupil response; sequence of lesson not rigid; teacher had difficulty getting attention at times.

AFFECTIVE CLIMATE

(6) *Consideration.* High on affective-imaginative statements and encouraging; courteous; shows awareness of pupil needs. High pupil interest.

(7) *Warmth.* High on support, praise; gentle reproof; more directive; reads questions from book or blackboard (!)

(8) *Disapproval.* High on reproving and criticizing; speech pattern rated low; terminates lesson abruptly.

Figure 1. Brief Descriptions of Eight Dimensions Measured in Classrooms of Elementary Student Teachers

ers from high and low ones, we would have identified—in thought at least—3^8 or 6,561 different kinds of teachers!) This is why we feel justified in speaking of "measuring the complex classroom of today."

Which of the 256 is the "best"? Which is the "worst"? Is it any wonder we have trouble measuring a single dimension called teacher effectiveness or teacher competence? Is it reasonable to assume such a dimension even exists? Perhaps our task is not a measurement task at all but a classificatory one.

Perhaps the most interesting (not to say disturbing) finding regarding the changes in the student teachers over a 10-week period on the eight dimensions was that a good half of them were idiosyncratic—that is, if one teacher increased on a particular dimension, another teacher in the same room was just about as likely to decrease on the same dimension as to increase, while the third teacher in that room might stay the same. Since all three teachers would be undergoing experiences as nearly identical as we could manage, including having the same supervisor and cooperating teacher, the effects of those experiences seem highly unpredictable.[1]

Insofar as there were changes common to most teachers, we should note that the typical student teacher tended to have more Presence after student teaching than before. (This tendency was not statistically significant.) He also tended to become more Informative but no more Imaginative in his teaching (this was significant); and his pupils increased more in Activity than in Initiative (also significant). But none of his behaviors related to affective climate changed significantly.

On the basis of these findings, and within their limitations, we might predict that the average student teacher would increase in his control in the classroom both of "discipline" and of the content discussed, as well as in the amount of pupil activity generated (or permitted). He would not grow in sensitivity to or affection for pupils, however, nor in use of criticism or in willingness to let pupils initiate procedure.

One might speculate that these changes would have survival value for a teacher during his first months in the classroom—that the observed increases in classroom control, content covered, and pupil activity reflect a sensible decision to be sure one is in full charge of one's class before trying to do anything fancy (that is, imaginative).

What teacher educators might ask themselves is whether these are desirable changes—the ones we are trying to bring about in students at

[1] No evidence was found that systematic differences in supervisory feedback procedures introduced as part of the experiment, or the particular supervisors assigned to student teachers, had any consistent effect on the behavior of the student teachers.

this point in their development. If these are the objectives of supervision in student teaching, they are being attained. If not, then we need to reexamine what we are doing.

As far as I can determine, no teacher education institution ever collects this kind of data about what is happening to its student teachers. Do I sound unreasonable when I say that they seem to me to be operating blindly? I think I have heard some of them cursing this darkness; is it time to light a candle somewhere?

A Verbal Category System

The second study I want to describe was done with full-time first-year intern teachers in their own public school classrooms (Medley and Hill). This group was atypical in that it was made up of liberal arts college graduates whose professional training had been limited to what they had obtained in one summer session plus a weekly seminar operated concurrently with the internship they were serving.

The subjects were 70 secondary school intern teachers of science, mathematics, English, or social studies in junior or senior high schools. They were visited in their classrooms four times—twice in early February and twice in the following May. On each occasion, behavior was coded on an instrument called OScAR 4V (Observation Schedule and Record Number 4, Verbal) (Medley *et al.*, 1967). As in the earlier study, variations between events observed about a week apart were regarded as evidence of instability of behavior over time; and differences between behaviors recorded several months apart were, insofar as they exceeded these short-term variations, regarded as stable changes in behavior with experience.

OScAR 4V is a verbal category system which yields frequency counts of the occurrence of each of 42 different "events." Two basic kinds of events are recognized—*statements*, in which one person speaks, and *interchanges*, in which a teacher and a pupil interact. Statements are classified according to their content—affective, substantive, procedural, etc. Interchanges are classified twice: first, according to the type of pupil behavior elicited; second, according to how the teacher evaluates the behavior.

A factor analysis of the records of these 70 teachers again yielded eight orthogonal dimensions, which I would like to describe briefly.

The first dimension had to do with the relative amount of lecturing vs. interaction with students in a classroom: we called it *lecturing behavior.*

The second dimension was called *question source,* and measured the relative numbers of questions asked by the teacher and by the pupils. It is worth noting that, according to the weights on this key, the teacher whose pupils ask a lot of questions is himself reacting differently to them than the one whose pupils ask relatively few. The kind of pupil behavior that the teacher reinforces is strongly associated with the amount of pupil initiative shown. Apparently, accepting whatever the pupil says reinforces the behavior of volunteering a comment or question, while giving approval only when the content is correct or acceptable reinforces the behavior of volunteering *correct* information only. The pupils soon learn to discriminate the teacher interested in hearing whatever they have to say from the one who is interested only in having them say certain things; and to keep quiet in the latter instance unless they are sure they are right.

This is an illustration of the fact that, in addition to measuring dimensions of teacher behavior, this instrument also provides specific diagnostic information indicating how a teacher should behave in order to score high on a dimension. To be able to say something to a teacher like "accept pupil statements or questions rather than approve them" should be more helpful than merely telling him that he needs to have more pupil initiations in his classroom, and leaving it to him to figure out how.

The third dimension is called *question difficulty.* Teachers score high on this, mainly giving positive feedback on pupil responses—that is, indicating that a pupil's answer or statement is correct without praising it. A teacher is more likely to praise a pupil when he answers a difficult question correctly than when he answers an *easy* question correctly. Since a teacher gets a high score on this scale by giving positive feedback without praise, a high score seems to mean that the teacher is asking easy questions.[2] Hence we infer that the main variable here is the difficulty of the questions the teacher asks.

The fourth dimension has been named *question type* because it has to do mainly with the pattern of questioning behavior the teacher uses. The high-scoring teacher asks relatively few convergent questions (that is, questions designed to elicit one particular answer from the pupil) and relatively many divergent ones (ones to which a pupil has some choice of responses) and elaborating ones (ones which ask a pupil to comment on or extend a previous answer). Moreover, once a pupil has answered a question, the high-scoring teacher tends *not* to evaluate his answer, but rather to ask a student to evaluate it or discuss it further.

[2] Each of the teacher behaviors making up this third dimension is derived from a separate category of the OScAR 4V system.

In such a teacher's class, it behooves a pupil to attend to what others say, and to think about it, because he may have to react to it.

At the opposite end of this dimension is a teacher who asks mostly questions that have only one right answer (convergent), who tends to evaluate the answers himself (give positive or negative feedback) and then ask an unrelated question of someone else. Pupils in such a class need not listen to others, so long as they pay attention to the teacher. Teachers low on this dimension do not conduct discussions, but recitations or quick-fire drill sessions. In brief, the scale contrasts two sharply different questioning styles; or, one might say, discussion with recitation.

As one studies these four dimensions, he recognizes them as describing questioning strategies which are not necessarily good or bad in themselves so much as appropriate for different purposes and at different times. However, our data, which were obtained on different days during the same month, clearly indicated that any given teacher tends to use the same strategy consistently, day in and day out. Rather than varying their strategy to suit their purposes, most teachers seem to fall into a habit of prescribing the same medicine for all patients, regardless of their symptoms.

The remaining four dimensions relate more to management than to instruction. The first of these is called *rebuking behavior.* It seems to reflect how often a teacher rebukes a pupil (rather than how intense or hostile the rebukes may be); it may very well be a reflection of the amount of difficulty a teacher has—or thinks he has—in maintaining order.

The next dimension is unique in that it is based entirely on the frequency of one class of events: continuing pupil statements. A teacher's score on this dimension depends on how often a pupil who has just said something to the teacher gets a chance to say something more *before the teacher interrupts him.* We have called it *listening behavior,* because the high scoring teacher seems to be one who listens to what pupils say and is not so anxious to speak himself as the average teacher is.

Permissive behavior is based on events in two categories. A teacher scores high on this dimension by frequently asking a pupil what he would like to do, and low by frequently refusing to let a pupil do something he has asked permission to do.

Managing behavior reflects the amount of verbal behavior a teacher devotes to discussing procedures to be followed in the classroom and trying to see that they are followed.

This second set of four dimensions was also found to be highly stable from day to day—that is, matters of habit. In other words, the teacher who spends a lot of time managing his class on one day tends

to spend a lot of time doing so every day. We must conclude that how one of these teachers behaved was less related to the particular situation that existed on any given day than to some customary pattern of behavior the teacher had consciously or unconsciously adopted.

This conclusion is relevant to our main concern, which is with changes in teacher behavior over time—from February to May in this case; it suggests that what we are looking for is changes in habits rather than in ways of reacting to classroom contingencies.

Presumably, the first year of teaching—particularly for teachers like these, who had had no supervised student teaching beforehand—is a period in which a teacher tries, pretty much on the basis of trial and error, to discover and acquire those methods and techniques which work best for him. If there are any methods or techniques which most beginning teachers find useful, this fact should reveal itself in the form of changes uniform in the entire group. If such generally useful techniques can be identified, it might be a good idea to include training in their use as part of the preservice teacher education program. Teachers who learned these methods before they began to teach would start something like a year ahead of where they would otherwise.

It is unfortunate that circumstances prevented us from observing the teachers at the beginning of the year and thus assessing a full year of change. Instead, we had to begin our study in February, so we could measure changes during the last half of the year only. It is possible, therefore, that we missed many of the changes that took place over the year as a whole. However, we found that as a group the teachers were still changing in a number of ways even after the year was half gone.

Significant changes were found in the questioning behavior of the teachers: by the end of the year, the average teacher was using *fewer* questions than at the start (and therefore lecturing more); in addition, the average level of difficulty of the questions asked was higher. None of the teachers changed in regard to the relative number of questions asked by pupils, or in the type of questions the teacher asked. Interestingly enough, the *rank order* of the teachers did change on this last dimension. This means that some teachers decreased in the amount of recitation or drill they used, and increased in use of discussion or thought-producing questions. It also means that others changed in the opposite direction—more drill and less discussion. Since both types of teachers are presumably trying to improve, this suggests a need for supervisory guidance.

Could the changes observed in the average teacher be described as improvement? Did the group improve with experience? Looking at these data, I am inclined to doubt it. Yet it is hard to believe that the average

teacher gets worse with experience. If (on a common sense basis) we argue that the average teacher *must* have improved, the implication is that, at this level of competence, better teachers lecture more and ask fewer questions and more difficult ones than they tend to ask naturally.

When we look at changes on the last four dimensions, things are not much better. We find the average teacher becoming less permissive and spending more time on management as he gains in experience—that is, as he (presumably) develops greater skill. Neither the amount of rebuking behavior nor the teacher's willingness to keep quiet and listen to his pupils changes, however. This last dimension—listening—was one on which individual teachers did change significantly, but in no consistent direction. Some teachers listen more, some listen less, at the end of the year than they did in the middle. Which ones are "right"?

Although these two studies of change in teacher behavior were done with different instruments in different places and in very different groups of teachers, both of them seem to suggest at least two practical conclusions.

One is that when teachers first confront actual pupils in the classroom they tend to underestimate their pupils' capacity for dealing with content and to overestimate their own ability to handle pupils. This is suggested by the fact that the elementary student teachers shifted from imaginative to informative teaching and increased in presence, and by the fact that the secondary interns increased the amount of lecturing they did and the difficulty of the questions they asked, and became more concerned with managing their class and less permissive. All of this is consistent with the notion that the first teaching experiences produce "reality shock" in teachers.

The second conclusion I would draw from these two studies has to do with the need these teachers have for precise data about their own behavior if they are to get control over the changes that take place. On two of what I would consider to be potentially the most important dimensions measured—question type and listening behavior—there was no uniform change. And yet it is clear that significant changes were taking place on both of them—half of the teachers going one way, half the other. Are these teachers aware of these changes? And are they getting any feedback on the effects of the changes—on whether as they change their behavior things get better or worse in their classrooms?

If measurements like those obtained in the second study had been available to these teachers, together with the sage counsel of mature supervisors, the blind trial and error that they s[illegible] to have used could have been replaced by systematic experimentat[illegible]hat would almost certainly have led to greater improvement in the [illegible]age teacher's skill.

Perhaps I am overinterpreting these data a bit. Certainly the substantive findings should be regarded as highly tentative. The important thing we learned from this study is that, despite the complexities of classroom phenomena, it is possible to measure improvement in at least some aspects of what goes on with sufficient precision to detect important changes when they occur. The use of a modern observation instrument makes it possible, then, for the supervisor to separate his function of feeding back information about teacher performance from his role as an advisor or evaluator. Teacher and supervisor together can look at the teacher's behavior objectively, decide what changes should be attempted, and, most important, determine objectively whether the proposed changes actually came about.

References

D. M. Medley. "Measuring Changes in Student-Teacher Behavior: A Second Look." Paper read at the annual meeting of the Association for Student Teaching in Chicago, Illinois, on February 20, 1964.

D. M. Medley. "The Language of Teacher Behavior: Communicating the Results of Structured Observations to Teachers." *Journal of Teacher Education* 22 (2): 157-65; Summer 1971.

D. M. Medley and R. A. Hill. "A Comparison of Two Techniques for Analyzing Classroom Behaviors." In: R. C. Burkhart, editor. *The Assessment Revolution: New Viewpoints for Teacher Evaluation.* Albany: New York State Education Department, n.d. Also: D. M. Medley and R. A. Hill. "Dimensions of Classroom Behavior Measured by Two Systems of Interaction Analysis." *Educational Leadership* 26 (8): 821-24; May 1969.

D. M. Medley, J. T. Impellitteri, and Lou H. Smith. *Coding Teachers' Verbal Behavior in the Classroom: A Manual for Users of OScAR 4V.* Reprinted in: Anita Simon and E. G. Boyer, editors. *Mirrors for Behavior: An Anthology of Classroom Observation Instruments.* Philadelphia: Research for Better Schools, Inc., 1967.

H. Schueler, M. J. Gold, and H. E. Mitzel. *The Use of Television for Improving Teacher Training, Phase I: Improvement of Student Teaching.* USOE Grant No. 730035. New York: Hunter College of the City University of New York, 1962.

J. Withall. "Development of a Technique for the Measurement of Socio-Emotional Climate in the Classroom." *Journal of Experimental Education* 17: 347-61; 1949.

6 • The Nonverbal Realities of Classroom Life

Charles M. Galloway

THE SIGNIFICANCE of nonverbal communication has long been recognized. Teachers now are confident that facial expressions, body movements, postures, and gestures do make a difference. Teachers also realize that furniture arrangements, uses of time, and travel patterns (movement to and from students) influence the tenor of classroom life. Students soon learn the meaning of teacher expressions. The eyes are dead giveaways. Glances and eye contact express support, disapproval, or neutrality. Specific gestures signify a class period is ended, an explanation is requested, or student behavior is disruptive.

Neither teachers nor students have been instructed in the meaning of these events. No teacher ever deliberately teaches such a curriculum. Nonverbal communication occurs as a practical matter and its significance is best understood at a common sense level. Teachers and students have long testified about the importance of nonverbal cues and their consequences; yet theoreticians and researchers have failed to investigate their influences on school life. Why has this been so?

The neglect has been fostered by several factors. The foremost reason has been the eagerness of researchers to study the verbal contents and patterns of classroom discourse. Educators are enamored with the economy and eloquence of verbal influence. To tell students what to do and how to do it has been supported historically as the *sine qua non* of teaching. Prior to the past 20 years, educators and researchers knew little of the precise influence of verbalization on student learning. It was believed that verbal contacts between teachers and students represented basic variables for research. Nonverbal behaviors were assumed to be consistent with verbal behaviors, and the actual influence of nonverbal cues was believed to correspond with verbal interaction. In other words, a valid sampling of verbal behavior was assumed to be an adequate sampling of nonverbal influence. Such an assumption, however, had little or no support from behavioral scientists.

The most overwhelming difficulty faced by investigators in any field has been determining which methods of analysis reveal the meanings of nonverbal information. No current method can claim to have a dictionary of definitions which describe such meanings. The profound research problem in nonverbal behavior is what is meant by meaning. Dictionaries are available to provide the meanings of words. References are made to dictionaries to discover word definitions, but word meanings are never complete until we understand their usage in context.

Words can be used to communicate almost anything. Verbal literacy is actually connected to our ability to understand and to use words, and dictionaries are enormously helpful. But the test of meaning relies upon the precise way a word is used and how it is responded to. Nonverbal cues and body languages suffer from a similar disadvantage, but the handicap is even greater because no dictionary of behavioral signs and signals with their definitions exists for handy reference. The meanings of nonverbal behavior are learned during human contacts, and no assurances can be given that one's working dictionary is valid and reliable.

To know that feelings can be conveyed through touch, facial expression, tone of voice, posture, rate of speech, body movement, etc., provides no assurance that one can detect when and how a feeling is communicated. Untrained adults and children easily infer that they are liked or accepted but may be unable to identify the bases for the inference. When a distinction is made between verbal (words) and nonverbal information (intonation, tone, stress, length, and frequency of pauses), each mode of expression (verbal or nonverbal) may not convey the same feeling. Sarcasm is easily recognized when a contradiction exists between verbal and vocal information. Usually the verbal message is positive while the vocal information is negative. When someone calls a person "honey" in a nasty tone of voice, two pieces of information are conveyed simultaneously, and the nonverbal information carries the heavier load of the meaning. Similarly, it is possible to say "I hate you" in a way that conveys an opposite intent.

Nonverbal Realities

Whenever human beings come into contact, a reality exists that is understood and shared without words. This is the fundamental assumption that undergirds the significance of nonverbal communication. People everywhere bear testimony to the assumption that nonverbal influences are recognized and understood. Since teachers and students engage in continual communicative contacts, it is reasonable to assume that nonverbal relationships exist.

Theoretical arguments have been promulgated by many scholars suggesting why nonverbal phenomena are significant to human relationships. Hall (1959), Birdwhistell (1970), Goffman (1959), Ruesch and Kees (1956), and Davitz (1964), to name a few, have provided imaginative explanations and descriptions of nonverbal realities. Perhaps the most adequate rationale and set of assumptions have been provided by Ekman and Friesen (1968).

Nonverbal behavior can be viewed as a relationship language. Silent cues signal a change or provide continuity for any interpersonal relationships. These cues, whether by face, eyes, or gesture, can be the primary means of expressing attitudes of intimacy, aloofness, concern, or indifference. Teacher attitudes can be inferred from the way a teacher looks at a student or looks to avoid him. Not only do special nonverbal cues appear to exist between a teacher and some students implying favorable relationships, but the very absence of these cues can be noticed between the same teacher and other students. Although differing teacher-student relationships can be quite evident on these nonverbal terms, little or no conversation occurs regarding this reality.

A second assumption, generally shared by psychologists, is that nonverbal behaviors are the primary vehicles for expressing emotion. Behaviors convey hate, fear, anger, anxiety, and other emotionalities. Feelings of pleasure or distrust can be transmitted by teacher or student. Although teachers may state their feelings in verbal forms, the existence of nonverbal signs can belie and contradict verbal utterances. Students often wonder whether a correspondence exists between what a teacher feels and what he says. Words may fail to be persuasive carriers of feeling since nonverbal behaviors are often more convincing.

Another assumption is that nonverbal cues function as qualifiers in the form of metacommunicative messages to indicate how verbal statements ought to be understood. For instance, a student at his desk may signify verbally that he is working but simultaneously act out a nonverbal performance that he is busy, believing that this kind of behavior is more convincing. While he may actually be working at his assigned task, much of his energy is spent in looking as though he is working. Often, a teacher will lack a certain firmness in his voice when remonstrating students to stop talking, causing students to surmise that it is okay to continue their conversation. Conversely, a smile, frown, or gesture can accompany a verbal request which makes the direction of the intended meaning very clear.

An assumption shared by behavioral scientists in several fields and strongly supported by psychiatrists is that nonverbal behavior provides a leakage channel which is difficult to control or to censor (Ekman and

Friesen, 1969a). In simple language, this means that nonverbal behavior is more likely to reveal true emotions and feelings and is less likely to be deceptive. Nonverbal behaviors betray one's feelings, whereas verbal communications are easier disguises for expressing feelings.

It is well known that most people are unaware of their body language and the feelings they convey to others. In ordinary circumstances one has no feedback available regarding the leakages of feelings that occur in body language. Verbal language offers the marvelous facility of providing immediate feedback, because a person can hear himself talk. But one is tempted to infer that others grasp the meaning of his verbal statements to the same extent that he understands the meaning of his own information. Whether information comes in the form of verbal or nonverbal messages, it is essential to obtain feedback and to recognize that leakages and misunderstandings can constitute the message.

A difficulty in monitoring one's nonverbal messages is that little feedback is available because a person cannot see himself. Others may comment on what someone says or how he says it, but little information is shared regarding body movement and expression. Our culture lacks a ready language for discussing nonverbal cues, and people are hesitant to discuss how others act to their faces. Students have long delighted in discussing among themselves the behavioral idiosyncracies of teachers, but rarely will they discuss them with the teacher himself.

We can assume that we are much less aware of our nonverbal behavior than our verbal. Goffman (1959) presents another view on this matter. He suggests that nonverbal behaviors can be managed to achieve a desired effect. His view emphasizes the idea that people in everyday life take on roles for the express purpose of achieving proper impressions. This does not mean, however, that impression management is easy. Everyone is not successful in achieving effects that are in his best interest. Despite the successes of behavioral management, which can be associated with courtroom lawyers, diplomats, used car salesmen, and others, nonverbal cues are less manageable and often more revealing than verbal information.

A final assumption about nonverbal behavior implies that learned patterns of body language are associated with what it means to be a teacher or student in school. Certain specified behavioral cues and responses are learned by teachers and students in their role-taking activities in classrooms. Teachers throughout this culture have been observed in the act of snapping their fingers to get attention, holding a finger to their lips to achieve silence, folding their arms to signify disapproval, staring directly at students to convey negative reinforcement, and pointing at students to give directions. These signs and signals

are well understood by students, and any observer can see the results.

Students also acquire behavioral cues necessary to their role as school-goers. They can be observed as looking as if they are listening, as appearing busy at work with their academic assignments, and as head-nodders who appear to understand teacher explanations and instructions. Students learn very early in school to raise their hands to be recognized, and they soon discover that hand-raising strategies are in their best interests. Body cues among teachers and students provide the means for influence when words would probably fail to be as effective. Many nonverbal behaviors are common to the performance of what it means to teach and to go to schools.

Why should it be necessary to say that nonverbal behavior provides unique information apart from verbal information? What is the significance of body language to classroom interaction and school life? Information seekers, whether they be teachers or students, will always search for extra data when they are not satisfied with verbal information alone. This condition of being discontent with the narrow range of verbal information and of relying on nonverbal data occurs when teachers or students are (a) unwilling or incapable of verbalizing information, (b) unapproachable to obtain information, or (c) uncertain about what is said verbally. In effect, body language speaks loudly when verbal information is missing or in doubt.

Problems in Studying Nonverbal Behavior

Although nonverbal behavior is a rich source of information, tough-minded researchers recognize the research difficulties. Problems continue to plague the unwary who believe that data are easily obtainable. It is all too clear that nonverbal studies are difficult to design. Measurable units of behavior are not readily available, and precise analytic methods have not been devised. Many nonverbal cues that appear in classrooms are elusive and ephemeral. Observers find data collection to be confounding and laborious. The very motivations that lead the researcher to observe the complexities of the classroom serve to be the most elusive kinds of data to identify and measure. When looking at nonverbal interactions between teacher and student, an observer is reminded again and again that human contact is difficult to analyze. Unless rigorous precautions are taken, a researcher will lack adequate measures of reliability and validity, and he will be uncertain of the usefulness of his information. Accuracy and fidelity are the historic problems which have confronted researchers in all fields of behavioral analysis.

Many of the early experimental studies of this century tested

whether observers and judges could accurately identify the emotions of subjects when specified emotions were expressed nonverbally. Stimuli were usually provided by photographs of posed expressions. Much of this work led to inconclusive results (Bruner and Tagiuri, 1954). One factor that precluded accurate judgments by observers was the absence of context. Missing a definition of the situation and an understanding of the context in which the expression occurred, observers were inconsistent in their judgments. Another factor in these early studies which prevented accurate estimates of emotion resulted from a reliance on posed expressions by actors. Furthermore, many of these posed emotions appeared unnatural to observers. From these early studies it was learned that an understanding of contextual information and the appearance of natural behaviors were necessary to studies of nonverbal behavior.

Three serious questions confront the researcher when he chooses to study and analyze nonverbal data: (a) when to look, (b) what to look for, and (c) how to observe. But no problem has been more difficult than the question of deciding what observational unit to use. The failure to develop useful categories has handicapped the study of nonverbal behavior. Observational categories developed by educational researchers have been too broad and too vague in their definition. Descriptions such as supportive, disapproving, positive, negative, attentive, and responsible are representative of these broad categories. Another limitation associated with choosing behavioral units has been the question of how long or short a unit of observation should be. Three seconds? Thirty seconds? Three minutes? A photograph? A frame of film? A naturally occurring unit of unspecified duration? Do you observe nonverbal behavior during speech, or do you reserve nonverbal observations for periods of silence? All nonverbal phenomena cannot be observed at once. An observer must make choices about when, what, and how to observe. There is an obvious difference between watching a teacher move from desk to desk and looking for eye contact and a mutual glance. In one context, movement among pupils may be a significant act; but in another situation, a glance carries a heavy loading of influence.

The research challenge facing students of nonverbal behavior is the collection of data which demonstrate that nonverbal cues provide crucial information unobtainable from studies of verbal behavior. Most of the basic research data in teacher behavior and student activity come from verbal and vocal behavior as revealed by typescripts and tape recordings. In theory and in the exploratory studies already achieved, it is clear that nonverbal behavior is a rich source of information that can be observed with profit. The need for data and evidence on the distinctive kinds of information that nonverbal behavior yields is a

necessary next step. When are nonverbal behaviors truly redundant with verbal messages? Under what conditions? How do they differ? What are the advantages of nonverbal behavior studies over analyses of verbal behavior?

Recent Research Approaches

When researchers turn to the study of nonverbal language, they become convinced that what human beings express with their body movements can be more informative than what they say. Anthropologists, sociologists, psychologists, and educationists have all addressed themselves to the reality of nonverbal influence. Each discipline brings a unique kind of explanation and insight to this emerging field of study.

Anthropologists are interested in cross-cultural studies of gesture and movement, looking for differences and similarities in body language. Their particular interest centers on how nonverbal languages are learned and what unique forms they take in cultural expression. In anthropology, studies of nonverbal language are based on analyses of cultural behavior. Anthropologists, such as Hall and Birdwhistell, view the term communication as synonymous with culture.

The work in sociology has been theoretical and explanatory, rather than empirical. Goffman (1959) reflects the bias of sociology when he states that human identity is a product of role performance. His work has been an analysis of the expressions a person *gives off* (nonverbal), rather than the expressions one *gives* (verbal). His speculative accounts of nonverbal influence have emphasized consistently the behavioral attributes of human contact. Unlike the psychologists, who choose to observe finite and specific modes of nonverbal expression, Goffman's analyses reflect a molar and general view. He writes of impression management, expression games, strategic interactions, discrepant roles, and territorial regions. He is interested in how the body codes or role management facilitates interaction with human contact.

Psychologists observe specific facial expressions and molecular body movements and are moving toward amassing a large amount of data that have implications for studies of psychotherapy and personality (Ekman and Friesen, 1969b). Human motivation and emotionality are the provinces of the psychologist, and studies of nonverbal behavior reflect this interest. Indeed, the "Freudian slip" is nothing more than extra information which reveals something that otherwise would not be available to the therapist. A recent work by Ekman and associates at the Langley Porter Institute on nonverbal behavior has begun to provide

data that can be extremely useful to psychiatrists and therapists (Ekman and others, 1971).

In the past score of years, educational researchers have been creating and developing numerous observational instruments that describe classroom interaction. Similar to the spirit of instrument development for verbal behavior, researchers have recently been busy constructing nonverbal observational schemes.

In Galloway's initial study of nonverbal communication in classroom situations, he attempted to develop an observational system to describe the consequences of nonverbal acts (Galloway, 1962). Bound by his pedagogical interest in the effects of teacher behavior on subsequent student behavior, he created observational categories that had broad rather than specific definitions. In a critical sense, his categories were not nonverbal definitions. The category system was composed of the following: (a) supportive, (b) helping, (c) attentive, (d) pro forma, (e) inattentive, (f) unresponsive, and (g) disapproving. The first three categories reflected teacher behavior which encouraged communicative contacts, and the latter three categories restricted teacher-student interaction.

The purpose of Galloway's study was to determine whether a reliable observational procedure could be developed. He succeeded in creating an observational instrument, but the data were not any more illuminating than the evidence which followed from Hughes' categories of controlling, teacher imposition, facilitating, positive affectivity, and negative affectivity; or Anderson's general claims that a difference exists between teacher behaviors which are dominative or integrative; or Flanders' distinction between direct and indirect teacher behaviors.

In a later extension of Flanders' initial category system, which describes verbal interaction, Galloway subscripted additional categories on each of the Flanders categories. By grafting on categories, Galloway attempted to describe teacher nonverbal behavior which accompanied verbal activity. Again, however, these additional categories were not strictly nonverbal in character. Pure nonverbal categories usually relate to facial activity, body movement, or gesture. Galloway's categories had pedagogical referents which implied teaching and learning consequences.

By extending Galloway's initial category system, Victoria (1970) was successful in developing a typology of nonverbal gestural behavior which was exhibited by student teachers in art. Observations of these teachers were made during specified contexts of task setting, demonstration, and evaluation.

The advantage of using category systems for observation is obvious. Their use implies economy of effort and their utility affords an

abbreviated version of note taking. But the data provided by category schemes can be limited in value, and the short-hand advantage can preclude observations of behavior which are not included in the definitions of categories. A promising approach for observing behaviors, which has received limited use, is the recording of nonverbal acts in narrative descriptions. Jackson's *Life in Classrooms* (1968) and Smith and Geoffrey's *Complexities of an Urban Classroom* (1968) provide excellent illustrations of the method of observing classroom phenomena by the use of written descriptions of behavior.

In each instance, taking the role of participant observer, these researchers wrote narrative accounts of what they saw and understood about classroom activity. Such an approach has heuristic advantages and is not limited to the deductive limitations of predefined categories. Indeed, categories of interest can emanate from the data when the researcher has taken an inductive approach rather than a deductive one. Writing narrative descriptions of nonverbal behavior enables the observer to note significant behaviors not otherwise included by previously defined categories.

The theoretical developments and the research results of the past several years are more suggestive than definitive. No ready-made schemes for training potential teachers or for use with teachers through in-service programs are currently available. Much of the work on nonverbal behavior proceeds piecemeal, with each trainer or researcher creating the emphases which he deems most desirable. Whether the focus is on noting teacher nonverbal behaviors or detecting the significance of silent student behaviors, the teacher educator enhances the study of pedagogy when studies of nonverbal cues are included in the curriculum. But the true import of nonverbal behavior for teacher education would emphasize both teacher and student nonverbal cues. Analyzing the influences and effects of nonverbal information from either source has significance for better understanding the nature of teaching and learning.

The nonverbal realities of classroom life reflect different classes of data which can prove useful to the practitioner and to the researcher. Nonverbal cues provide information to both participants and observers. This implicit information represents the hidden realities and the unspoken meanings of what is to be understood. Information is always available, whether it be in the form of furniture arrangements, duration of class periods, facial expressions, gestures, or vocal intonations and inflections. All of this occurs whether the teacher and student are aware of its meaning or not. Nonverbal information is always available in some form. But information is not always communication.

A communicative act occurs when a teacher or student intentionally

attempts to send a message. An intention to communicate differs from the sheer availability of information. A sender must deliberately attempt to convey a message for it to be considered as communication. Nonverbal communication implies that information is available at a level of awareness and that a conscious effort is made to transmit a message. When information is available to a teacher or to a student that is not intentionally communicated, then these data are merely informative. When a teacher or student acts to influence the other, these actions are recognized as interactive. Interaction is marked by the distinction of achieving influence on the perception and behavior of another. In effect, classroom events can be informative, communicative, and interactive. Nonverbal signs, signals, and events can provide information regarding the realities of classroom life; they can occur as intentional efforts to communicate expectations and instructions; and they can appear as moves to influence perceptions and behaviors.

Teachers provide information for students; they intentionally communicate to students; and they interact with students. These data are made available in verbal and nonverbal ways. And, of course, students inform, communicate, and interact with teachers. The profound implication for teacher education rests with our need to collect and analyze the significance of nonverbal acts and events during teaching and learning.

References

Ray L. Birdwhistell. *Kinesics and Context.* Philadelphia: University of Pennsylvania Press, 1970.

Jerome Bruner and R. Tagiuri. "The Perception of People." In: Gardner Lindzey *et al. Handbook of Social Psychology.* Cambridge, Massachusetts: Addison-Wesley Publishing Company, 1954. pp. 634-54.

Joel R. Davitz. *The Communication of Emotional Meaning.* New York: McGraw-Hill Book Company, 1964.

Paul Ekman, P. Ellsworth, and W. V. Friesen. *The Face and Emotion.* New York: Pergamon Press, 1971.

Paul Ekman and W. V. Friesen. "Head and Body Cues in the Judgment of Emotion: A Reformulation." *Perceptual and Motor Skills* 24: 711-24; 1967.

Paul Ekman and W. V. Friesen. "Nonverbal Behavior in Psychotherapy Research." *Research in Psychotherapy* 3: 179-216; 1968.

Paul Ekman and W. V. Friesen. "Nonverbal Leakage and Clues to Deception." *Psychiatry* 32 (1): 88-105; February 1969a.

Paul Ekman and W. V. Friesen. "The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding." *Semiotica* 1 (1): 48-98; 1969b.

Charles Galloway. "An Exploratory Study of Observational Procedures for Determining Teacher Nonverbal Communication." Unpublished doctoral dissertation, University of Florida, 1962.

Erving Goffman. *The Presentation of Self in Everyday Living.* New York: Doubleday and Company, 1959.

Edward T. Hall. *The Silent Language.* New York: Doubleday and Company, 1959.

Philip Jackson. *Life in Classrooms.* New York: Holt, Rinehart and Winston, Inc., 1968.

Jurgan Ruesch and Eldon Kees. *Nonverbal Communication.* Berkeley: University of California Press, 1956.

Louis Smith and William Geoffrey. *The Complexities of an Urban Classroom.* New York: Holt, Rinehart and Winston, Inc., 1968.

James J. Victoria. "An Investigation of Nonverbal Behavior of Student Teachers." Unpublished doctoral dissertation, Pennsylvania State University, 1970.

7 • Use of the Flanders Interaction Analysis System

Larry S. Bowen

THE FLANDERS system of interaction analysis (IA) has received a great amount of attention and has been used in two ways by groups of educators during the past decade. First, this 10-category, socioemotional-oriented instrument has been used extensively in *researching* verbal behavior in classrooms. The findings of that research have been reported widely (1).

A second use of the instrument has been in *teacher education and training* programs at both pre- and in-service levels. It is to this concern that the following remarks are directed. The primary objective is to describe an in-service program that employed Flanders' system in a way that did then, and does now, make some ethical sense to me for the contemporary supervision enterprise. For the supervisor, IA can serve an important function: as a reality-testing instrument for the teacher to "see" his own and the student's verbal behavior in the classroom. The teacher can gain self-insight into his real behavior and can use his IA data to experiment with varying patterns of verbal behavior. Before describing the in-service program, however, I shall present a quick sketch of the system of IA.

The System

Flanders' technique consists of 10 categories of verbal behavior which can be identified by the observer from either a taped or a live lesson. The first seven categories are "teacher talk," these being divided into direct and indirect influence. Categories 8 and 9 are "student talk," and category 10 is "silence or confusion." A summary of the 10 categories is seen in Figure 1.[1]

[1] Reprinted by special permission from: Ned Flanders. *Analyzing Teaching Behavior.* Reading, Massachusetts: Addison-Wesley Publishing Company, 1970. p. 34.

Teacher Talk	Response	1. *Accepts feeling.* Accepts and clarifies an attitude or the feeling tone of a pupil in a nonthreatening manner. Feelings may be positive or negative. Predicting and recalling feelings are included. 2. *Praises or encourages.* Praises or encourages pupil action or behavior. Jokes that release tension, but not at the expense of another individual; nodding head, and saying "Um hm?" or "go on" are included. 3. *Accepts or uses ideas of pupils.* Clarifying, building, or developing ideas suggested by a pupil. Teacher extensions of pupil ideas are included but as the teacher brings more of his own ideas into play, shift to category five.
		4. *Asks questions.* Asking a question about content or procedure, based on teacher ideas, with the intent that a pupil will answer.
	Initiation	5. *Lecturing.* Giving facts or opinions about content or procedures; expressing *his own* ideas, giving *his own* explanation, or citing an authority other than a pupil. 6. *Giving directions.* Directions, commands, or orders to which a pupil is expected to comply. 7. *Criticizing or justifying authority.* Statements intended to change pupil behavior from nonacceptable to acceptable pattern; bawling someone out; stating why the teacher is doing what he is doing; extreme self-reference.
Pupil Talk	Response	8. *Pupil-talk—response.* Talk by pupils in response to teacher. Teacher initiates the contact or solicits pupil statement or structures the situation. Freedom to express own ideas is limited.
	Initiation	9. *Pupil-talk—initiation.* Talk by pupils which they initiate. Expressing own ideas; initiating a new topic; freedom to develop opinions and a line of thought, like asking thoughtful questions; going beyond the existing structure.
Silence		10. *Silence or confusion.* Pauses, short periods of silence and periods of confusion in which communication cannot be understood by the observer.

Note: There is *no* scale implied by these numbers. Each number is classificatory; it designates a particular kind of communication event. To write these numbers down during observation is to enumerate, not to judge a position on a scale.

Figure 1. Flanders' Interaction Analysis Categories (FIAC)

In using the technique, the following two-step procedure is utilized: (a) every three seconds the observer tallies the observed category; thus, for a 20-minute lesson the observer will have approximately 400 tallies; (b) the observer then places the numbers on a matrix in sequential pairs in such a way that each number is entered twice, once as the first and once as the second number in the pair.

The rows of the matrix (Figure 2) represent the first number in the pair, and the columns, the second. As an example, consider the following six numbers tallied: 4, 8, 8, 3, 5, 5. The first sequence pair, 4-8, is tallied in the cell that is located at the intersection of row 4 and column 8. Following are the sequence pairs 8-8, the next in 8-3, the next in 3-5, and the last in 5-5. These five pairs represent about 15 seconds of talk. The matrix reveals the placement of the five pairs discussed above, and from analysis of the tallies taken from a five-minute or more lesson, the teacher can determine the degree of emphasis in such areas as: (a) content; (b) student ideas; (c) teacher authority; (d) control; (e) pace of lesson; and (f) acceptance and praise.

The In-Service Program

In the summer of 1967, the El Dorado County Office of Education in Placerville, California, coordinated an in-service program for 28

		Second Event										
		1	2	3	4	5	6	7	8	9	10	Total
First Event	1											
	2											
	3											
	4											
	5											
	6											
	7											
	8											
	9											
	10											
Total												

Figure 2. Interaction Analysis Matrix

elementary and intermediate school teachers from that county. Cooperating institutions were the University of California at Davis and the 12 school districts in the county. The four-week institute was conducted at the Buckeye Elementary School in Shingle Springs, California. The institute was held daily for 19 days from 8:30 a.m. to 12:30 p.m. A four-week demonstration school for 75 children in grades 1 through 6 was held in conjunction with the institute. Students from the elementary school were involved in the in-service program. The teachers received extension course credit from the University of California at Davis.

A major objective of the institute was to train participants in the application of the interaction analysis system. Program designers assumed that teachers who were skilled in use of the technique were in a better position to analyze their own verbal teaching behavior, identify their strengths and weaknesses, and revise their verbal behavior according to newly established goals. The role of the leaders of the institute was one of accepting and cultivating that assumption.

The teachers received instruction in the meaning and use of Flanders' interaction analysis tool. They gained knowledge and use of some new ideas and instructional materials in science education, the content vehicle for the microteaching experiences they were to have. Video tape recorders were used for obtaining feedback. The seven teaching activities of motivating, informing, leading a discussion, planning, counseling, disciplining, and evaluating (2) constituted the skills of the teaching act.

Participants cooperatively planned science lessons that were individually taught to groups of four to six children from the demonstration school. The intended verbal patterns were coded on matrices; the real verbal interaction revealed on the five- to ten-minute videotaped lessons was likewise coded.

So What?

The process which *followed* the coding of the observed lessons was a crucial part of that institute. Supervision typically breaks down or lives during that process. "To judge or to help?" That became the question which was, and must now be, confronted. Does the supervisor coerce the teacher into doing what the former considers "good," or does he insist upon teacher self-description/evaluation?

The contemporary supervisor must demonstrate, in his work with teachers, the importance of self-evaluation. He must resist the temptation of providing answers to questions that do not yet exist. He can do that by responding to the teacher's spoken and nonverbal pleas for help by

not telling him what to do, but rather by helping him clarify what he *wants* to do. Instructional decision making can be placed on the teacher by helping him to do and get what he himself decides is of value, and *not* by telling the teacher what he ought to do. The supervisor can help project alternative means with which the teacher can reach his target (model). In other words, the supervisor primarily must be interested in pulling from the teacher that which the teacher wants to do, serving as a resource for selection of alternative means, and presenting tests of reality to the teacher where and when needed.

Reality testing is found in IA. The supervisor shows he cares; he supports the objective decided upon; but he holds the teacher and himself to standards that can be stated, reexamined, modified, and adhered to. Supervisors in the El Dorado Institute, then, were the teachers themselves. The instructors served as supervisors, too, and emphasized, through their behavior, the self-evaluation model described here. Crucial to the success of the institute was an understanding and belief in Carl Rogers' helping relationship: psychological maturity. Although that is outside the scope of this booklet, I suggest that supervisors should give equal consideration to Rogers' point of view in addition to acquiring kn wledge of observational techniques.

The greater the psychological maturity of the supervisor, the more he will view and use such observational systems for nonoppressive ends. As a consequence of experiences in seeing himself, the teacher becomes a more autonomous person. The greater his autonomy, the greater value he 'll see in that condition for *his* teaching of children. To evaluate the outcome at the end of the institute was most difficult. Attitude change was the only measurement outcome selected for study (3). Informal observation in institute teachers' classes the following year yielded exciting impressions of teacher and student behavior which suggested definite growth in that respect. In practice, then, the use of interaction analysis can serve as an important tool to describe the realities in the classroom and help teachers compare congruencies between reality and intention.

References

1. E. J. Amidon and J. B. Hough. *Interaction Analysis: Research Theory and Application.* Reading, Massachusetts: Addison-Wesley Publishing Company, 1967.

2. E. J. Amidon and Elizabeth Hunter. *Improving Teaching: The Analysis of Classroom Verbal Interaction.* New York: Holt, Rinehart and Winston, Inc., 1967.

3. Larry S. Bowen. "Influence of Feedback as a Way of Changing Teacher Attitudes." Unpublished master's thesis, California State University, Hayward, 1968.

8 • Self-Evaluation Through Video Tape Recordings

I. V. Ahnell and Horace C. Hawn

LITTLE IS known of the value of videotaped recordings in teacher education. Their use in teacher education programs has, however, increased dramatically in an attempt to modify behavior. This increase could be attributed to two advantages cited in a Hunter College research project: video tapes are readily available for repeated viewings, and they can be viewed by more than one individual simultaneously (Schueler and Gold, 1964).

Increased portability and reasonable costs of videotaping equipment have also had an impact upon its use. In a review of microteaching in 1967 Gustafson reported that many institutions around the country were using microteaching in various forms. He stated that, "All of those whom the writer identified as using microteaching were convinced of its merit, but offered no evidence to support these conclusions" (Gustafson, 1967).

There has been extensive use of videotaping to provide feedback for behavior modification. Some studies (for example, Aubertine, 1967; Olivero, 1964) have indicated some possibilities in this area. Others (for example, Bedics, 1970; Ward, 1970; and Kantz, 1970) did not acknowledge the use of videotaping as a superior method of providing feedback. Kerber found little modification of the self-concept among selected student teachers (Kerber, 1967).

Extensive utilization of videotaping in microteaching was made at Brigham Young University. Clark Webb and others reported that, in a sample of 730 students, 75 percent said it was "very true" or "somewhat true" that microteaching "changed my self-image as a teacher." Ninety-seven percent responded that microteaching "indicated to me areas where I can make improvement" (Webb, 1968).

Students seemed to find videotaping helpful in analyzing their teaching techniques and procedures. It would appear that if behavior change were to occur through the use of video tape recordings it might

well be in the area of self-concept and self-evaluation. Such behavior changes could help teachers attain professional maturity.

Video tape recordings (VTR) have been used to provide feedback for selected student teachers at the University of Georgia over the past three years. Reports from student teachers, supervising teachers, and college supervisors who used VTR were overwhelmingly positive. The writers shared concern over the growing interest in videotaped recordings during the student teaching program without evidence to indicate they brought about any significant behavior modification of prospective teachers. A longitudinal study was designed and begun in the fall of 1970 to study this problem.

It was assumed that realistic self-evaluation was necessary for a teacher to improve his teaching performance. An important objective of teacher training programs is to help students develop ways of evaluating their effectiveness in the classroom. Improved teaching performance can come only by a change in teaching behavior. A change in teaching behavior will come only as the teacher sees the need for a change. Such a need could be realized from self-evaluation.

Working within a perceptual psychology frame of reference, it was logical to assume that an individual teacher could change his self-perceptions and, if this occurred, he could evaluate his teaching behavior with a view toward improving it. It was also assumed that these changes would relate to his feelings of personal adequacy and his open- or closed-mindedness.

Methodology

The sample for the first year of the study, 1970-71, was 68 female subjects. Thirty-five in the experimental group were placed in one student teaching center and 33 in the control group were placed in another. A 10-minute teaching tape was made of each subject during the first week of the student teaching quarter. Each experimental subject evaluated her own tape and the tapes of all other experimental subjects by viewing and recording her evaluations on a VTR Observation Index. Control subjects did the same with themselves and other control subjects. Both experimental and control groups used this same procedure, viewing their initial tapes as a post-measure at the end of their student teaching quarter. All subjects were pre- and post-tested with a battery of tests including the Rokeach Dogmatism Scale, the Tennessee Self-Concept Scale, and a semantic differential measure of 10 selected concepts.

The treatment for the experimental group consisted of videotaping each subject in conjunction with the college supervisor's observations

six times during the quarter. The videotaped recording was provided as immediate feedback to the student teacher by her reviewing it with the college supervisor. The control group did not utilize VTR during the quarter.

Analyses of variance, covariance, and correlation coefficients were used to determine statistical significance among the data obtained.

Preliminary Conclusions

Data gathered after one year of the study have provided no conclusive evidence, but they have given direction for the second year of research. They have suggested ways by which the student might benefit from viewing himself on video tape.

It would seem that viewing one's self over a relatively short period of time does not change his belief-disbelief system as measured by the Rokeach Dogmatism Scale. However, further analysis of these data will be made to explore the possible relationship between open- or closed-mindedness and self-evaluation.

Several questions arose as a result of the lower self-concept attained by the experimental group after the use of video tape recording during their student teaching. Did the viewing of themselves cause self-concept to regress? Was the experimental group more realistic in the concept of self in the post-test? Did the supervisory techniques vary enough to cause this difference? Further research is taking place to inquire into these questions.

It was assumed that self-evaluation would require self-criticism. Because both self-evaluation and self-criticism did not vary between experimental and control groups, it would appear that the treatment of the experimental group had no effect.

For the purposes of this study, reality in self-evaluation was defined as the mean of the evaluation by group peers. It seemed necessary to establish some criterion of realism for self-evaluation. It was assumed that realistic self-evaluation would be achieved by a positive correlation of self-evaluation and peer evaluation scores. It was found that all subjects in this study evaluated themselves realistically according to this assumption. It was evident that the treatment did not affect this variable. This suggested that prospective teachers in both groups were capable of realistic self-appraisal, a necessary requirement in the process of becoming an effective teacher.

The experimental group saw in the concept "Evaluating Teaching by Videotaping" an evaluative factor, whereas the control group saw in it a potency and activity factor. This indicated that the treatment

emphasized the value of video tape feedback as an evaluative technique. The control group appeared to see only the peripheral characteristics of videotaped classroom interactions. Because the control group was not subjected to the treatment, it can well be understood that they might have viewed videotaping as a potentially strong and busy activity.

In conclusion, it could be that the sparse results in video tape self-evaluation may be ascribed to the constraining limitations of the measuring instruments or inadequate research design. The continuation of this research with modified measuring instruments is under way. It seems that video tape recording as a self-evaluative technique must have powerful potential if the assumptions of perceptual psychology are valid. And these assumptions warrant further research.

References

Horace E. Aubertine. "The Use of Micro-Teaching in the Process of Training Clinical Supervisors." Paper read at AERA conference, New York City, 1967.

Richard A. Bedics. "A Study of Self-Evaluations of Student Teachers Through the Use of Video-Tape." Unpublished doctoral dissertation, University of Alabama, 1970.

Kent L. Gustafson. "Microteaching." *Program and Papers of the Second Workshop for Directors and College Supervisors of Student Teaching.* East Lansing: Mic State University, 1967. p. 67.

illiam D. Kantz. "Effectiveness of Videotape Feedback and Systematic Observation Training in Modifying the Verbal Behavior of Student Teachers." Unpublished doctoral dissertation, Pennsylvania State University, 1970.

James E. Kerber. "Study of Selected Student Teachers and Modifications of Their Self-Concept Through Use of Portable Videotape Recordings." Unpublished doctoral dissertation, Wayne State University, 1967.

James Lee Olivero. "Video Recordings as a Substitute for Live Observations in Teacher Education." Unpublished doctoral dissertation, Stanford University, 1964.

Herbert Schueler and Milton J. Gold. "Video Recordings of Student Teachers—A Report of the Hunter College Research Project Evaluating the Use of Kinescopes in Preparing Student Teachers." *Journal of Teacher Education* 15: 359; December 1964.

Phillip M. Ward. "The Use of the Portable Videotape Recorder in Helping In-Service Teachers Self-Evaluate Their Teaching Behavior." Unpublished doctoral dissertation, University of California, Berkeley, 1970.

Clark Webb and others. "Description of a Large Scale Micro Teaching Program." Provo, Utah: Brigham Young University, 1968.

9 • Content Analysis of Mathematics Instructional Materials

William C. Lowry

A FRUITFUL approach to making content analyses of instructional materials in mathematics appears in some models which have been developed to provide organizational frameworks for specifying the desirable outcomes of mathematics learning. In the main such models take into account two dimensions of mathematics learning, the content dimension and the learner levels-of-behavior dimension. The model, or table of specifications, then, often takes the form of a "content" by "levels-of-behavior" matrix. A number of different classifications of content have been used, depending largely on the grade level of the students involved and whether the content classifications are meant to be descriptive of what most current programs are doing or whether they are meant to be prescriptive of what a desirable program should be like.

Perhaps the best known and most widely used classification scheme for levels of behavior in the cognitive domain is that developed by Bloom (1956). The six levels described and illustrated in this taxonomy are knowledge, comprehension, application, analysis, synthesis, and evaluation. These levels are ordered, in that the behavior at any one level in the list is held to be more cognitively complex than behaviors at any of the levels below it. Thus knowledge behaviors are the least and evaluation behaviors the most cognitively complex. The list is also hierarchical in that the behavior at any one of the levels may, and usually does, involve behaviors from one or more of the preceding levels. One could then use three large content areas of school mathematics—arithmetic, algebra, and geometry—to construct a "content" by "levels-of-behavior" matrix.

Each of the six levels of behavior can be subdivided. For example, Bloom's taxonomy has three such subdivisions under knowledge: knowledge of specifics, knowledge of ways and means of dealing with specifics, and knowledge of the universals and abstractions in a field. Each of these subdivisions is further divided. And, of course, the three major

categories of content can be subdivided many times. For example, major subdivisions of secondary school algebra include such topics as algebraic expressions, equations and inequalities and their solutions, and relations and functions. One could, then, conceive of an n x m matrix where n and m are fairly large numbers.

It is not necessary, of course, to stick strictly with Bloom's taxonomy in classifying levels of behavior. Several investigators have found variations of this taxonomy more suitable to their respective fields. Wilson (1971, pp. 646-47) has provided a carefully developed and thoroughly illustrated table of specifications for evaluating learning in secondary mathematics. The reader should note that Wilson's classifications are ordered and hierarchically arranged in a like manner to Bloom's taxonomy. Wilson also includes a classification of affective behaviors in his table of specifications.

The model can be used, of course, in the development of instructional materials. For nearly all teachers and supervisors, however, the use of the model with respect to instructional materials would be in a different way: (a) to aid in the selection of published materials for a course or an entire program, (b) to aid in selecting from the already adopted materials those parts which best conform to the model, or (c) to assess the relevance of class materials and assignments to the developing competencies of individual pupils.

If one were concerned with content only in the traditional sense, the levels-of-behavior dimension could be bypassed. Thus the mathematical topics, in whatever fashion they are treated in the instructional materials, could simply be checked against the items in the content dimension of the model to see if the emphases given to the various topics in the instructional materials are commensurate with the emphases suggested for those topics in the model. Yet such an approach to content analysis, in this writer's opinion, would be incomplete in light of the spirit of contemporary thinking about the objectives of mathematics learning. Educators are interested not only in what content a student encounters but also at what intellectual or cognitive level he is able to work with what content. It is at this point that systematic observation and classification are necessary in relation to a student's successful and unsuccessful responses to various problems representative of different cells in the method.

The final part of this paper adapts Wilson's model for certain topics of a seventh grade mathematics program and shows how it can be used to make a content analysis of instructional materials, in this instance textbooks. Many contemporary mathematics textbooks are developed in a heuristic style, with the authors attempting to carry on a conversation

with the reader. Even so, this part of the textbook, insofar as the student reader is concerned, is still largely expository in nature. The student seeks explanations of how to do things by reading the author's discussion. The exposition and the various lists for student work can be checked for content coverage. Yet it is probably in the exercise lists, problem lists, and—in some texts—exploratory or discovery activities that the various levels of behavior expected of the student show up best.

There are exceptions, of course. The student who reads carefully probably gains at least as much exposure from the exposition as he does from the student work sections to Wilson's category B.5, the ability to follow a line of reasoning. The cognitive levels of behavior and some of the content classifications under number systems described by Wilson are used in the model in Figure 1; and two items typical of those found in problem, exercise, or exploratory-discovery lists in current seventh grade mathematics textbooks are used to illustrate the use of the model in making an analysis of instructional materials.

Example 1. 6402 (base seven) + $15\overline{36}$ (base seven) = ?

If this exercise comes after the student has done considerable computation in other bases, including base seven, the level of behavior is A.3, ability to carry out algorithms, in the content category 1.9, numeration systems. On the other hand, if the student has had a bare introduction to computation in, say, base five only, finding this sum in base seven numerals may require a level of behavior on his part as high as D.1, ability to solve nonroutine problems. This illustrates that in order to classify a given item properly while making an analysis of instructional materials, one must know what preceded that item in the materials; and even then it has to be assumed that all students who will encounter the item will have about the same proficiency in what preceded. Student proficiency records based on analysis of their previous work permit an even more personal classification to be made; this illustrates the place of observational data in content analysis.

Example 2. If x is a positive whole number and x is added to both the numerator and the denominator of 2/3, will the resulting number be greater than, less than, or equal to 2/3? Or will it depend on what positive whole number x is?

For most seventh grade students, this would be an item classified as D.2, ability to discover relationships (What relationship is there between the two fractions under these conditions?), in the content category 1.3, rational numbers. To say this is a D.2 item is to say that the ability to discover relationships is the highest level of cognitive behavior

Content	Cognitive Behaviors																					
	A.0 Computation			B.0 Comprehension						C.0 Application				D.0 Analysis								
	A.1 Knowledge of specific facts	A.2 Knowledge of terminology	A.3 Ability to carry out algorithms	B.1 Knowledge of concepts	B.2 Knowledge of principles, rules, and generalizations	B.3 Knowledge of mathematical structure	B.4 Ability to transform problem elements from one mode to another	B.5 Ability to follow a line of reasoning	B.6 Ability to read and interpret a problem	C.1 Ability to solve routine problems	C.2 Ability to make comparisons	C.3 Ability to analyze data	C.4 Ability to recognize patterns, isomorphisms, and symmetries	D.1 Ability to solve nonroutin roblems	D.2 Ability to discover relationships	D.3 Ability to construct proofs	D.4 Ability to criticize proofs	D.5 Ability to formulate and validate generalizations
1.0 Number systems																						
1.1 Whole numbers																						
1.2 Integers																						
1.3 Rational numbers																						
1.4 Real numbers																						
1.6 Finite number systems																						
1.8 Probability																						
1.9 Numeration systems																						

Adapted from: B. S. Bloom, J. T. Hastings, and G. F. Madaus, editors. *Handbook on Formative and Summative Evaluation of Student Learning*. Copyright © 1971 by McGraw-Hill, Inc. Used with permission of McGraw-Hill Book Co. pp. 646-47.

Figure 1. Model for Evaluation Learning in Secondary School Mathematics

required to answer the question. Within the hierarchical arrangement of these behaviors, the student makes use of several lower level behaviors. Examples are A.2, knowledge of terminology ("positive whole number," "numerator," "denominator"), and B.6, ability to read and interpret a problem. He may even arrive at his answer mainly by use of C.2; he tries several specific values for x, compares the two fractions in each case, and arrives at the relationship. Again, closely questioning him about or observation of his solution attempts is necessary to make an accurate assessment of the cognitive processes he actually used.

In order to use the model to help select a textbook or to help select the material one will use from the textbook, decisions must be made as to the amount of emphasis to be placed on each of the "content"/"level-of-behavior" cells in the matrix. The analysis of the textbook gives a picture of how well the textbook conforms to the emphases desired. The textbook can then be adopted or rejected, or material can be selected from it to conform more nearly to the model.

An accusation, often subjectively arrived at, of instructional materials, tests, and even teachers is that while they espouse the higher levels of cognitive behavior as being very desirable outcomes of mathematics learning, they in fact emphasize the lower level behaviors. The use of models such as those discussed here to make analyses of instructional materials provides a more objective way to determine the relative emphasis placed by the materials on the various levels of behavior expected of the learner in dealing with the content. Also, if individual observations are made of pupils' responses to the various types of material representing the model, content analysis becomes a dynamic means for assessing the materials of instruction in relation to teaching expectancies and learning outcomes.

References

B. S. Bloom, editor. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain.* New York: Longmans, Green and Co., Inc., 1956.

D. R. Krathwohl, B. S. Bloom, and B. B. Masia. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook II: Affective Domain.* New York: David McKay Co., Inc., 1964.

James W. Wilson. "Evaluation of Learning in Secondary School Mathematics." Chapter 19 in: B. S. Bloom, J. T. Hastings, and G. F. Madaus. *Handbook on Formative and Summative Evaluation of Student Learning.* New York: McGraw-Hill Book Company, 1971.

10 • Analyzing the Content of English Instruction: A Point of View

Joseph E. Strzepek

I WOULD like to propose some ideas which supervisors might use in analyzing the content of instruction in English classes. Analysis of the content of classroom instruction can serve many purposes: it can enable the supervisors to help teachers clarify their objectives, and analyze and evaluate both their instructional techniques and student behaviors. It is a process that ought to include, particularly in the field of English instruction, the participation of students.

Traditionally, it has been expected that besides knowing observational and supervisory strategies, the English supervisor needs to know how the subject matters of English language and literature, and the component skills of thinking, speaking, and writing about (and listening to) those subject matters, can be categorized and organized developmentally in a K-12 curriculum.

In the *Handbook on Formative and Summative Evaluation of Student Learning* (Bloom *et al.*, 1971), three chapters present exhaustive charts diagramming the content of behaviors and skills in English in terms of cognitive and affective objectives: (a) Moore and Kennedy dealt with elementary language arts; (b) Purves presented literature; and (c) Foley covered writing. Although they were aimed at helping teachers write test items to produce desired behaviors, these chapters should be very useful references to anyone attempting to classify or describe the constant and stable factors in English classes: the curriculum guide, the unit or lesson plans, assignments, even a teacher's classroom questions or directions.

Also, education literature has assumed that the supervisor needs to assess how individual teachers perceive the subject matter in relation to the specific characteristics of their students' ages, abilities, and backgrounds. Those perceptions will probably determine the teacher's objectives—the intended effects of the English instruction. The supervisor should confer with the teacher; analyze his plans and tests; and ask the

teacher to articulate his objectives and give his impression of how the students responded to the instructions.

In most cases, content analysis in English stops at this point. English supervisors and teachers may be moving toward the viewpoint, espoused by Moffett (1968) and many British teachers (Squire and Applebee, 1969), that the content of English is not merely a restricted canon of literature or dramatic works, linguistic theories, or rhetorical principles, but almost everything that happens in the English classroom, specifically the original verbal and dramatic behaviors of the students. Often these behaviors will be responses to published literary works, but just as often they will be self-responses: their own talking, acting, and writing. From this viewpoint, the content of English classes becomes, potentially, more interactive, personal, creative, unpredictable, and more difficult to analyze for the supervisor-observer.

Moffett urges that transcripts and recordings of class discussion and student writing be made. Certainly students are playing larger roles in determining the content of English instruction; they request particular courses in elective curricula and increasingly undertake independent studies and projects that they contract and negotiate with their teachers.

With these new activities and attitudes in English, it becomes increasingly important that teachers and students carefully record the processes and products of instruction: the in-class happenings, the papers and projects produced by students, tests, and the evaluations of the students. Only then will teachers and supervisors be able to analyze and evaluate the original and serendipitous learning along with classroom behaviors that are dull, commonplace, even destructive to learning. How else can a school which has made a curriculum change evaluate the results of the change, say from a traditional to an elective curriculum?

If English teachers are to move toward evaluation by comment instead of the gross letter grading system, they need to learn the observational skills of the reporter, the ability to cite detail of a clinician, the analytical skill of a critic, and the artistry of good authors. Consider what Herbert Kohl wrote in *36 Children* and *Teaching the Unteachable;* John Holt, in *Why Children Fail;* Charles Silberman, in *Crisis in the Classroom;* what James Herndon said in *The Way It Spozed To Be;* and what Daniel Foder and Morton Shaevitz presented in *Hooked on Books.* These men, while they were teaching or observing, kept careful account of the students' behaviors, especially what they said and wrote. By merely presenting dramatic accounts or examples of those behaviors, these writers have probably had more influence on teaching English in the past 10 years than all previous researchers on observational methodology and instrumentation.

It's a Miracle [1]

Why was I created in this
World of frustration and damnation?

It's a miracle—

The addict on Lennox Ave. that thrives
On the $5 fix,
Who steals and kills
Just for kicks.

It's a miracle—

How man can construct and then destroy.

It's a miracle—

The White, Black, Yellow and Red
Inhabitants of this city and
This damned World.

The lush on Saturday, the Christian
On Sunday and back to the lush on
Monday.
How can these creatures exist;
The sacrilegious, the religious
fanatic, the drug addict?
It's miraculous I tell you,
It's miraculous.

As I walk the streets of Manhattan
I sniff the contaminated air
Which was meant to smell like satin.

Where trees used to exist.
Gigantic structures pop out.
Wherever you go.

Oh
I wish it wasn't so.

Heaven only knows
It's a miracle.

Why must I exist in this city
Or in this World of hate, sin
and destruction?
It's miraculous I tell you, It's down right miraculous.

—Otto Grant, age 15

[1] Herbert Kohl. *Teaching the Unteachable.* New York: The New York Review Book Series, 1967. Reprinted with permission from *The New York Review of Books.* Copyright © 1967 by Herbert Kohl.

Dittoed class magazines are graphic examples of the students' work, their recorded talk and ideas. A major part of Kohl's *Teaching the Unteachable* (students in ghetto New York schools) is his reprinting of the student writing in *What Is Happening*. This magazine had an impact on the teachers, the students, and the school board. One poem from the magazine that Kohl reprinted appears on p. 72.

There are many ways to record classroom events, with audio and video recordings, transcriptions of classroom dialog, saving copies of student papers and projects, anecdotal observations by supervisors, teachers, and students. Whatever the behavior recorded, I suggest that the following kinds of questions be raised as a means of analyzing the content of the event. An event could be an assignment, a question, a project, paper test, part of a class, a whole class, a unit.

- Who is speaking to whom for what purpose, and how valid is that purpose?
- What is (was) the effect of the event on the participants?
- How was it understood/responded to?
- How was the event organized; what form did it take?
- What was the meaning of it?

I suggest that these questions and others should be discussed by udents, teachers, and supervisors. The answers to these questions may reveal as much about the value premises of the respondents as they will about the event. Those value premises ought to be articulated so that the people (teachers, school boards, parents, students) for whom the analysis is made know how to interpret it. Methods used by the content analyst in English instruction are not unlike those used by semanticists and literary critics. The basis of these methods is a systematic pattern of inquiry designed to produce a description, interpretation, and evaluation of the content.

The English supervisor has advantages the literary critic usually does not. He can and should interview teachers and students to get their perceptions of a classroom happening; authors and characters are not so accessible. Nowadays, neither the English supervisor nor the teacher is the sole dominant analyst. In fact, students increasingly will be trained to record and analyze the content of their English classrooms. Such training may or may not include training or practice in the Flanders or OScAR procedures, the social psychology of groups. Formal observation instruments, though invaluable for research and training of supervisors, are, however, impractical and unwieldy for the everyday use of a supervisor who is an English specialist. Recording the actual content of the classroom experience and analyzing it to discover its various

meanings seem to be necessary to accomplish one of the primary goals of learning: to be able to construct an accurate picture of one's experience, reflect on it, and make sense of it, so as to generate new experiences that one would like to have.

References

Benjamin S. Bloom, J. Thomas Hastings, and George F. Madaus, editors. *Handbook on Formative and Summative Evaluation of Student Learning.* New York: McGraw-Hill Book Company, 1971.

Daniel N. Foder and Morton H. Shaevitz. *Hooked on Books.* Berkeley, California: Berkeley Publishing Corp., 1966.

James Herndon. *The Way It Spozed To Be.* New York. Simon & Schuster, 1968.

John Holt. *Why Children Fail.* New York: Pitman Publishing Corporation, 1964.

Herbert Kohl. *Teaching the Unteachable.* New York: The New York Review Book Series, 1967.

Herbert Kohl. *36 Children.* New York: New American Library, Inc., 1968.

James Moffett. *Student-Centered Language Arts Curriculum, Grades K-13: A Handbook for Teachers.* New York: Houghton Mifflin Company, 1968.

Charles E. Silberman. *Crisis in the Classroom.* New York: Random House, Inc., 1970.

James R. Squire and Roger K. Applebee. *Teaching English in the United Kingdom.* New York: Appleton-Century-Crofts, 1969.

11 • Instrumentation in Monitoring and Recording Human Behavior

Harold R. Strang and James R. George, III

AUTOMATED PROCEDURES for monitoring and recording human behavior have received considerable use. In addition to the widespread application of instrumentation in experimental settings, educational environments have increasingly relied upon both computer-based systems and electromechanical teaching machines not only to deliver instructional materials but also to accumulate ongoing measures of performance. For example, Strang (1971) developed an automated tutor capable of keeping ongoing accounts of initial accuracy, accuracy after tutoring, and time spent both in studying before answering and in using tutoring. Even the recording of standardized test performance has been successfully automated.

Turning from school-oriented academic performance to more generalized behaviors, technological advances and unrestricted importation have resulted in a great proliferation of reliable tape recorders, thus allowing teachers and any other would-be observers to collect permanent auditory records of behavior in almost any natural setting.

The tape recorder's usefulness as a scientific tool, however, has been limited by tape length and the awareness of observed subjects that they were being recorded. Bernal, Gibson, Williams, and Pesses (1971) developed a device which partially overcame these shortcomings. The coupling of a standard tape recorder with an inexpensive recycling timer allowed for periodic unannounced samplings of verbal behavior throughout an extended period of time. Observer-recorded reliability was .86.

Also relating to taped observation, Herbert and Swayze (1964) developed a technique called wireless observation which allowed for the synchronizing of taped auditory records of classroom activities with ongoing observer records of nonaudible events. The simple electronic system allowing for such recording consisted of two wireless microphones, an FM receiver, and a stereophonic tape recorder.

With the recent advent of inexpensive, dependable, and portable video tape recording (VTR) units, the classroom observer has yet another

valuable recording aid. Videotaping has been helpful in checking observer reliability in controlled observation areas such as specimen description and event sampling. Videotaping has not only allowed for multiple viewing of observation periods; but also, because of its applicability both as a model and a purveyor of self-feedback, has aided in the training of human observers.

With the expanding application of behavior modification in classroom settings has come an increased interest in facilitating time sampling, a technique common both to behavior modification and to observational study. At first glance it would seem that automating this technique, which simply consists of viewing behavioral frequency or duration over time, would present little technical difficulty. Indeed, this is the case when each subject can be equipped with a relatively inexpensive, unobtrusive device for transducing and keeping a quantitative record of the behavior under observation. Pedometers can be used to record walking as proximity switches to record in-seat behavior. Of a more complex nature, Schulmann and Reisman (1959) found that a standard self-winding wristwatch could easily be modified to record a measure of general hyperactivity in subjects. Goldman (1961) recorded body movements through a wrist-mounted device utilizing the Doppler principle. In numerous situations, of course, total instrumentation would prove disruptive or economically unfeasible.

Several applications of instrumentation to time sampling have not totally replaced the human observer but have greatly facilitated his gathering of observational data. Lindsley (1968) found a commercial golf score counter helpful in gathering behavioral rates. Worthy (1968) developed a time-regulated auditory device which facilitated a human observer's recording by prompting him at preset time intervals.

Turning to wider applications, Eleftheri-, Shoudt, and Strang (1972), while transducing behaviors through a hand switch closure, were able to accumulate automatically an accurate time-frequency record on the out-of-seat behavior of an entire class.

In a more recent study an attempt was made to effect a fully-automated observation-recording system in a natural classroom setting. As in the Eleftherios, Shoudt, and Strang study, the system was designed to monitor a single collective class signal. The behavior "observed" was class noise. The setting was a rural elementary school. The students were 24 second graders.

The "automated observer" consisted of two components. One, a monitor-sender unit, was secured in a central location on the classroom's ceiling. This small box transduced classroom sound of a preset amplitude to a radio signal. Transduction was accomplished by a voice operated

relay (VOR) wired in tandem to a radio signal generator The second element of the automated observer was a receiving unit housed in a two foot by four foot case and placed in a conspicuous location in the room. This unit received and then transduced the radio signals emanating from the sender to a relay switch closure which in turn operated the recording equipment. One clock accumulated total session times; another accumulated above-threshold noise time. A digital counter recorded the number of above-threshold outbursts. Such recordings afforded both rate and duration measures of the observed behavior.

The automated observer was used during a specific time of the day. This was a half-hour session directly following lunch when students were to put their heads down to rest at their desks while the teacher left the room for a planning period. Prior to observation, teacher reports indicated that the class noise level during this time was so excessive that neighboring classes were disturbed.

After the equipment had been installed in the classroom, the students guessed variously that it was a TV set, a camera, and a robot. For about three weeks after installation the teacher manually calibrated the ceiling monitor-sender so that a radio signal was produced beginning at a noise level deemed as "excessive" and "disturbing" her.

While the utility of the system as an observer-recorder seemed apparent, a second phase of the study was initiated whereby the system's possible use as an effector of behavioral change in the classroom was tested. Th·· ., on the twelfth day of operation a student feedback circuit on the receiving-recording unit was activated. This circuit consisted of a display of five jeweled lights arranged vertically on a panel within the view of all students. The students were informed that they could earn a light if they were not noisy during their rest period. In order to collect a light, the noise level in the room had to stay below threshold for a 30-second interval. For each "quiet" interval, another light in the column was earned, and if the students remained below threshold for six consecutive intervals, they collectively earned a gold coin displayed on a large plexiglass panel above the column. If an outburst occurred, however, during the climb to the gold coin, all the vertical lights were lost and the class collectively had to start over again. Those symbolic gold coins that had been earned were not lost.

After the students had been familiarized with the rules, they dubbed the apparatus "The Magic Ear."

Whatever one's philosophical position regarding the use of external behavioral controls, it should be noted that "The Magic Ear" effected a substantial reduction in some very undesirable and disturbing behaviors. During the 10 days of student feedback, excessive noise occurred only

half as often as before. Atypical of much intervention, such changes re: ed from the application of consistent visual feedback rather than the doling out of tangible extrinsic rewards.

Recent modifications have been directed toward animating the apparatus so that the symbolic coin awards have been replaced by a human-like face whose features systematically light up if the noise level is not excessive. Pilot results indicate that students respond well to such feedback. After all, who wants to stop a face from smiling!

In conclusion, then, instrumentation's inherent operative consistency not only influences intervention effectiveness, but also validates the use of such technology in recording applications. Equipment described in this paper can accumulate and store almost flawless records of a variety of student behaviors. In viewing the current scope of such applications, one must further realize that instrumentation technology is barely emerging from its infancy.

References

M. Bernal, D. Gib , D. Williams, and D. A. Pesses. "A Device for Automatic Audio-Tape Recording: *ournal of Applied Behavior Analysis* 4 (2): 151-56; 1971.

C. Eleftherios, J oudt, and H. R. Strang. "The Game Machine: A Technological Approach to Classroom Control." *Journal of School Psychology* 10 (1): 55-60; 1972.

J. A. Goldman. "A Look at Human Measurements in Industry." In: L. E. Slater, editor. *Proceedings of the Interdisciplinary Clinic on the Instrumentation Requirements for Psychophysiological Research.* New York: Foundation for Instrumentation Education and Research, 1961.

J. Herbert and J. Swayze. "Wireless Observation." Horace Mann-Lincoln Institute of School Experimentation pamphlet. New York: Columbia University Bureau of Publications, 1964.

O. R. Lindsley. "A Reliable Wrist Counter for Recording Behavior Rates." *Journal of Applied Behavior Analysis* 1: 77-78; 1968.

J. L. Schulmann and J. M. Reisman. "An Objective Measure of Hyperactivity." *American Journal of Mental Deficiency* 64: 455-56; 1959.

H. R. Strang. "An Automated Approach to Remedial Reading." Paper presented at the meeting of the American Educational Research Association, New York, 1971.

R. Worthy. "A Miniature Portable Timer and Audible Signal-Generating Device." *Journal of Applied Behavior Analysis* 1 (2): 159-60; 1968.

12 • Observation in Supervisory Practice and School Research

Richard M. Brandt and Hugh V. Perkins, Jr.

WHAT DOES this rather diverse assortment of papers suggest for the practicing supervisor or curriculum specialist? What main threads weave them together to help the reader translate their meaning into solid educational improvements?

Obviously, the central thrust is toward the greater and more intelligent use of observation in supervisory practice and school research. But does not modern supervision theory tend to de-emphasize the traditional role of supervisors' making classroom observations and then giving teachers suggestions for improving their teaching? Would not the tools and systems proposed in this booklet merely increase the efficiency of supervisors' observations and thereby stimulate still greater threat and control over teachers?

While it is possible that supervisors would use observational tools to enhance their power over teachers, such an outcome is not necessarily the inevitable result. Almost any technique can be used for good or ill; but what use is actually made depends on one's purposes and concern for people rather than the inherent quality of the tool. There is, in the nature of observational methodology, nothing to suggest that its use will automatically build barriers in the relationship of supervisory or administrative personnel with classroom teachers.

Quite the contrary. Accurate and sharply focused observations may actually improve communication and respect between these groups. If teachers and supervisors truly trust each other and hold similar purposes in observational data gathering, much more useful information can be obtained than was usually available in the past. The sheer complexity of the school classroom defies an easy analysis. The potential of observational methodology is great for simplifying some of this complexity to the point at which it can be better understood by both teachers and supervisors.

The potential of observational methodology can be recognized as

one sifts out the main threads running through these papers. Some of these threads pertaining to the specific nature of the instruments are:

1. Observational data consist increasingly of systematically obtained low inference classifications of behavior according to various behavior typologies, that is, category systems. A judgment made by the observer is primarily qualitative, indicating which type of behavior out of a limited and carefully defined set of categories a particular action represents. Such judgments are usually quite objective, with different observers in high agreement as to the type of action coded.

2. When ratings are made, furthermore, the trait or behavior being judged is often specified rather precisely so that it is readily distinguishable from other traits or behaviors. Increasingly, separate ratings are made of different instances in which a trait or behavior is displayed rather than making single ratings to cover a large number of discrete manifestations. This procedure not only reduces the ambiguity of what actually is being rated but permits assessment of trait variability from one time or setting to another.

3. While anecdotal type records are still functional for providing relatively accurate event records which can serve as a basis for later discussion and hypothesis generation, their scientific quality is questionable. Too much happens to record everything, and the selection of what to include in and exclude from the record, as well as what events to record at all, is usually haphazard at best. The utility of videotaping, audio recording, or even using automated equipment for obtaining rather full, faithful records of complex happenings in sequence is clearly evident in several of the papers. Such relatively complete narrative records permit later coding of classroom happenings in considerable depth and on a great variety of dimensions.

4. While observations can be quite useful when focused on only a relatively narrow segment of the total classroom happenings, for example the types of questions teachers ask or the amount of noise in a room, a number of highly complex systems are becoming available for coding activity on several dimensions at once as it occurs. Medley's OScAR and PROSE schedules are illustrative of such systems.

5. Many of those important variables, which because of difficulties in their assessment were often disregarded in early classroom studies, are now included in these more sophisticated observation category systems. Videotaping likewise contributes to the assessment of such hard-to-measure variables as classroom climate and nonverbal interaction. With

tapes, complex action sequences can be played over and over so that these and other variables can be coded covering the same basic episodes.

For the practicing supervisor and classroom teacher, the value of these developments in observational technology lies not only in the improved quality of educational research from which educators derive and evaluate their notions about appropriate instructional practice, but also in the direct use of some of this technology to enhance the quality of observations they themselves make. Many of the observational tools described in this volume can be readily adapted to the specific purposes and concerns which teachers face daily as they attempt to understand how well certain instructional procedures work or, more generally, what actually does happen in their own classrooms throughout a school day.

Instructional leaders would seem to have a special responsibility for becoming acquainted with observational technology so they can help teachers recognize both their own instructional patterns and particular performance and response characteristics of individual pupils. Objective observational records, particularly of those aspects of school life which the teacher is concerned about and wishes to have studied, can serve as highly useful feedback devices. Undoubtedly, as Medley indicates, they are much more acceptable to a teacher than either the opinions or ratings which supervisors have provided so often in the past. With precise, accurate records of some of the activity a teacher is responsible for but too busy to assess accurately while he is teaching, a teacher and supervisor should have most fruitful discussions of what the data mean, whether too much of some kinds of behavior and too little of others have occurred, and what instructional strategies might produce more desirable behaviors.

Perhaps the most sweeping generalization to be derived from the classroom studies conducted so far is that there is no simple, successful teaching formula to cover all classes, subjects, and educational objectives. Instead the classroom has been shown to be a highly complex matrix of many important and interacting variables. Pupil background, interest, and ability factors; specific characteristics of instructional materials; the nature of assignments and activities; grouping patterns; the quality and timing of feedback regarding the correctness of pupil responses; evaluation criteria and teaching expectancies; and many other interacting variables together determine what is learned or not learned. What the teacher does is important but far from sufficient to guarantee successful learning. Many other factors are important as well. Successful formulas, furthermore, may apply to many individuals or situations but seldom

to all. Just as children learn in different ways, teachers teach differently and often with equal success.

Even content analysis has been shown to be not merely a static process of assessing the various emphases stressed in particular instructional materials in relation to knowledge or ability taxonomies, but also appraisal of the levels of behavior expected of learners as they confront these materials. Assessment of these levels of behavior, furthermore, may vary from one learner to another, depending on the amount of previous exposure to similar materials and his own response patterns. A dynamic content analysis of this type would require some observational or conversational attempt to determine the precise levels of behavior employed.

What these observational tools permit, then, is a close scrutiny of one's own teaching activities concurrently with pupil behavior and response patterns as a prelude to discovered insight for teachers and greater learning for pupils. Self-fulfillment for teachers must come, at least in part, from knowing what one is accomplishing in relation to what is possible of accomplishment. A realistic understanding of classroom complexity and the nature of specific teaching functions within it would seem a basic condition for teacher self-fulfillment.

Obviously, to gather observational data, cooperative [illegible] activity is required, involving teachers, supervisors, and, as several papers indicate, perhaps students and aides also. The same kind of sharing and trust is necessary as that characterizing other supervisor-teacher activities. One person performs while the other observes and records. While both persons should be in agreement over what kinds of questions need to be answered by the data gathering, it is not always desirable for the performer to know specifically what or how behavior is being monitored until afterward, so as to minimize tendencies to change one's performance, often unconsciously, from whatever game plan is being followed. Afterward, both performer and observer can interpret the records together and attempt answers to the questions they posed earlier.

The relatively neutral quality of observational records should permit a close, nonthreatening relationship between the partners making up this diagnostic team, unless one (especially the supervisor) lets the other know that certain patterns are more acceptable than other patterns. The value judgments to be placed on various patterns of activity that have been monitored should come primarily from the performer. Indeed the observer may need to be especially supportive of the performer whatever patterns are observed, as self-scrutiny is not easy even within a trusting relationship.

One suggestion to help build both understanding of particular

observational tools and the partnership trust necessary for them to be used effectively is for supervisors and teachers to reverse customary roles, with the former conducting lessons while the latter observe and record. In going over records afterward, furthermore, supervisors should emphasize a diagnostic question-raising tone rather than the goodness or badness of particular teaching patterns while attempting to interpret what happened.

The interest in observation has come about in part to counterbalance (a) a very strong emphasis upon laboratory research in the behavioral sciences which underlie educational theory, and (b) an almost exclusive dependence on standardized tests, questionnaires, and poorly designed rating scales for evaluating school programs.

Educators have long recognized (a) the inappropriateness of applying laboratory findings directly to classroom situations without additional field testing, and (b) the shortcomings of even well-constructed measurement devices for assessing the full range of educational outcomes. The suggestion is not to disregard laboratory research totally nor to do away with testing programs, but to add observational methodology to our investigative activity in order to increase its scope and its applicability to real-life conditions and many of the critical dimensions of school life that have been neglected so long.

Obviously, as Raths so succinctly points out, observational methodology is not without its problems. Limitations exist in all measurement systems, including observational ones. Such limitations must be clearly recognized and reduced whenever possible. Overall, however, the observational tools described herein provide great promise for exciting investigative adventures ahead for educators ready to accept the challenge.

•Contributors

I. V. AHNELL, Assistant Professor of Education, University of Georgia, Athens

CHARLES W. BEEGLE, Associate Professor of Education, University of Virginia, Charlottesville

LESLEE J. BISHOP, Professor of Education, University of Georgia, Athens

LARRY S. BOWEN, Assistant Professor of Education, University of Virginia, Charlottesville

RICHARD M. BRANDT, Professor and Chairman, Department of Foundations of Education, University of Virginia, Charlottesville

ROBERT S. FLEMING, Professor of Education, Virginia Commonwealth University, Richmond

CHARLES M. GALLOWAY, Professor of Education, The Ohio State University, Columbus

JAMES R. GEORGE, III, Administrative Director, Year Round Project, Richmond School System, Virginia

HORACE C. HAWN, Assistant Professor of Education, University of Georgia, Athens

WILLIAM C. LOWRY, Professor of Education, University of Virginia, Charlottesville

DONALD M. MEDLEY, Professor of Education, University of Virginia, Charlottesville

HUGH V. PERKINS, JR., Professor of Education, University of Maryland, College Park

JAMES RATHS, Professor and Chairman, Department of Elementary Education, University of Illinois, Urbana

HAROLD G. SHANE, University Professor of Education, Indiana University, Bloomington, and President 1973-74, Association for Supervision and Curriculum Development, Washington, D.C.

HAROLD R. STRANG, Assistant Professor of Education, University of Virginia, Charlottesville

JOSEPH E. STRZEPEK, Assistant Professor of Education, University of Virginia, Charlottesville

ASCD Publications, Summer 1973

Yearbooks

Balance in the Curriculum (610-17274)	$5.00
Education for Peace: Focus on Mankind (17946)	$7.50
Evaluation as Feedback and Guide (610-17700)	$6.50
Fostering Mental Health in Our Schools (610-17256)	$4.00
Freedom, Bureaucracy, & Schooling (610-17508)	$6.50
Guidance in Curriculum (610-17266)	$5.00
Individualizing Instruction (610-17264)	$4.00
Leadership for Improving Instruction (610-17454)	$4.00
Learning and Mental Health in the School (610-17674)	$5.00
Learning and the Teacher (610-17270)	$4.50
Life Skills in School and Society (610-17786)	$5.50
New Insights and the Curriculum (610-17548)	$6.00
A New Look at Progressive Education (610-17812)	$8.00
Perceiving, Behaving, Becoming: A New Focus for Education (610-17278)	$5.00
Role of Supervisor and Curriculum Director (610-17624)	$5.00
To Nurture Humaneness: Commitment for the '70's (610-17810)	$6.00
Youth Education: Problems, Perspectives, Promises (610-17746)	$5.50

Books and Booklets

Bases for World Understanding and Cooperation: Suggestions for Teaching the Young Child (611-17834)	$1.00
Better Than Rating (611-17298)	$2.00
Beyond Jencks: The Myth of Equal Schooling (17928)	$2.00
The Changing Curriculum: Mathematics (611-17724)	$2.00
The Changing Curriculum: Modern Foreign Languages (611-17764)	$2.00
The Changing Curriculum: Science (611-17704)	$1.50
Changing Supervision for Changing Times (611-17802)	$2.00
Children's Social Learning (611-17326)	$1.75
Cooperative International Education (611-17347)	$1.50
Criteria for Theories of Instruction (611-17756)	$2.00
Curricular Concerns in a Revolutionary Era (611-1785)	$6.00
Curriculum Change: Direction and Process (611-17698)	$2.00
Curriculum Decisions ←→ Social Realities (611-17770)	$2.75
A Curriculum for Children (611-17790)	$3.00
Curriculum Materials 1973 (17922)	$2.00
Dare To Care / Dare To Act: Racism and Education (611-17850)	$2.00
Differentiated Staffing (17924)	$3.50
Discipline for Today's Children and Youth (611-17314)	$1.50
Early Childhood Education Today (611-17766)	$2.00
Educating the Children of the Poor (611-17752)	$2.00
Educating the Young People of the World (611-17506)	$2.50
Educational Accountability: Beyond Behavioral Objectives (611-17853)	$2.50
Elementary School Mathematics: A Guide to Current Research (611-17752)	$2.75
Elementary School Science: A Guide to Current Research (611-17726)	$2.25
Elementary School Social Studies: A Guide to Current Research (611-17384)	$2.75
Ethnic Modification of Curriculum (611-17832)	$1.00
Freeing Capacity To Learn (611-17322)	$2.00
Guidelines for Elementary Social Studies (611-17738)	$1.50
Human Variability and Learning (611-17332)	$2.00
The Humanities and the Curriculum (611-17708)	$2.00
Humanizing Education: The Person in the Process (611-17722)	$2.75
Humanizing the Secondary School (611-17780)	$2.75
Hunters Point Redeveloped—A Sixth-Grade Venture (611-17348)	$2.00
Improving Educational Assessment & An Inventory of Measures of Affective Behavior (611-17804)	$3.00
Influences in Curriculum Change (611-17730)	$2.25
Intellectual Development: Another Look (611-17618)	$1.75
International Dimension of Education (611-17816)	$2.25
Interpreting Language Arts Research for the Teacher (611-17846)	$4.00
The Junior High School We Saw (611-17604)	$2.00
Language and Meaning (611-17696)	$2.75
Learning More About Learning (611-17310)	$2.00
Linguistics and the Classroom Teacher (611-17720)	$2.75
A Man for Tomorrow's World (611-17838)	$2.25
New Dimensions in Learning (611-17336)	$2.00
Nurturing Individual Potential (611-17606)	$2..
Observational Methods in the Classroom (17948)	$3.50
On Early Learning: The Modifiability of Human Potential (611-17842)	$2.00
Open Schools for Children (17916)	$3.75
Personalized Supervision (611-17680)	$1.75
Removing Barriers to Humaneness in the High School (611-17848)	$2.50
The School of the Future—NOW (17920)	$3.75
Selecting New Aids to Teaching (611-17840)	$1.00
Social Studies Education Projects: An ASCD Index (611-17844)	$2.00
Strategy for Curriculum Change (611-17666)	$2.00
Student Unrest: Threat or Promise? (611-17818)	$2.75
Supervision: Emerging Profession (611-17796)	$5.00
Supervision in a New Key (17926)	$2.50
Supervision in Action (611-17346)	$1.50
Supervision: Perspectives and Propositions (611-17732)	$2.00
The Supervisor: Agent for Change in Teaching (611-17702)	$3.25
The Supervisor's Role in Negotiation (611-17798)	$1.00
Toward Professional Maturity (611-17740)	$1.50
The Unstudied Curriculum: Its Impact on Children (611-17820)	$2.75
What Are the Sources of the Curriculum? (611-17522)	$1.50
Child Growth Chart (618-17442) min. order 10 for	$2.00

Discounts on quantity orders of same title to single address: 10-49 copies, 10%; 50 or more copies, 15%. Make checks or money orders payable to ASCD. All orders must be prepaid except those on official purchase order forms. Shipping and handling charges will be added to billed purchase orders. **Please be sure to list the stock number of each publication, shown above in parentheses.**

Subscription to **Educational Leadership**—$8.00 a year. ASCD Membership dues: Regular (subscription and yearbook)—$20.00 a year; Comprehensive (includes subscription and yearbook plus other publications issued during period of the membership)—$30.00 a year.

Order from: **Association for Supervision and Curriculum Development**
1201 Sixteenth Street, N.W. Washington, D.C. 2003

END