CREDE
Center for Research on Education, Diversity & Excellence

Talking Leaves

Director's Corner

Under the Microscope

Classroom observation involves many participants, motives and purposes, and sometimes produces contradictory interpretations. The observation process not only measures, but affects the classroom–affects the teacher, perhaps the students, perhaps even some decision makers' future choices. Small wonder that good observers go carefully to the task, and that teachers watch them warily.

In this issue of *Talking Leaves*, we hope to remind all our readers (researchers, teachers, policy makers) of the complexity of the classroom observation process, so that mutual understanding can help us find common interests and common respect. For example, no teacher will ever cooperate willingly with a researcher who provides no prior orientation or later report on their observations; our teacher focus group report makes that very clear. Researchers often become understandably

In This Issue

Editor's Note: This issue of Talking Leaves *is devoted to classroom observation techniques used in CREDE research projects. In September 1999, CREDE researchers from around the U.S. gathered at an "Each Teach" conference in Santa Cruz, CA, where they shared their experiences and methods for observing teachers and students at work. The researchers discussed systems for collecting and analyzing data, using data as a tool for teacher professional development, and using observation findings to inform current and future research. Effective methods of classroom observation vary, but the central goal of improving education remains consistent across them all.*

CREDE Director Roland Tharp with Lilia Monzo and Gretchen Guiton at the Each Teach conference

Classroom and Schoolwide Observation
Rosemary C. Henze, ARC Associates

In our CREDE project, "Leading for Diversity" (*Talking Leaves 2*, 1), we are documenting how 21 school leaders addressed racial/ethnic conflicts and promoted more positive interethnic relations in their schools. As our research team collected data, we applied many of the practices and concepts of ethnography–a cluster of methods for gathering data, and analyzing, interpreting, and writing about a group of people who are considered a cultural unit in terms of day-to-day interactions. One of the most important ethnographic methods is observation.

Observations enabled us to document how school leaders promoted positive relations among different ethnic groups, and to describe variation in approaches within and across schools. Observation also helped us to confirm or disconfirm what was reported to us in interviews. Tying interviews with observations also illuminated the socially constructed nature of interethnic relations. The practices and approaches we observed in schools carried with them a history, belief

system, and assumptions about reality, which interviews helped us uncover. The combination of observations, interviews, and collection of documents provided one way to triangulate findings.

Our initial interviews with site leaders and others helped us identify the settings and practices that exemplified their approach to interethnic relations. These constituted our "working hypotheses," which we revised as we discovered other relevant settings. Examples of observable settings included classes designed to promote more positive relations or increase students' understanding of ethnic diversity, school 'family' or 'house' meetings, conflict resolution programs, leadership team meetings, cultural assemblies, district human relations committees, and parent involvement activities.

We developed specific guidelines for classroom and non-classroom observations. We focused on questions such as: What is the ethnic composition of those in the setting? How are students grouped? Who is leading or facilitating and how do they appear to exercise leadership? Are people talking about ethnicity/race/diversity? What are they saying? What stages or segments comprise the activity? Using the ethnographic technique of writing "thick descriptions," we took ample fieldnotes to allow someone else to visualize the setting.

For our fieldnotes, we separated observations from commentary or interpretive notes. We found that including diagrams to illustrate arrangements of people was helpful; as arrangements changed, we drew additional diagrams. In dialogue notes, we distinguished a real quote from an approximation. Patterns that emerged in these

"We have evidence indicating schools can make a difference in race relations…"

notes over time led us to assertions or claims about what is happening in a particular school or across schools.

Because our observation system is flexible, holistic, and open-ended, it is more difficult for researchers to learn to use than a more highly structured method. It can also have poor reliability. Even with common training, different researchers notice different things. Thus, we tried to visit sites in pairs whenever possible so that an entire case study was not dependent on one researcher.

A strength of ethnography, however, is its ability to improve study validity. Observation is highly con-textualized in the school's larger environment, and the clear separation of observations from interpretations helps maintain validity. By using an open-ended system, we were able to respond to a wide variety of observation opportunities, capturing more than would be possible with a highly structured system. When researchers disseminate their findings, they often use vignettes from their fieldnotes to illustrate the patterns they find. These vignettes help communicate to a larger audience, including non-researchers.

Observations are tools that help researchers apply findings to social issues. In our project, the social issue is whether schools can become a vehicle for improving race relations. People might argue that schools are uniquely unqualified to do this because of their history of segregation and continued use of sorting structures such as tracking. We have evidence indicating schools can make a difference in race relations, at least locally, and we hope to use our observations to support this point as we communicate study findings to a variety of audiences. 🍁

Read past issues of Talking Leaves online: www.cal.org/crede/

Observing School Restructuring in Multicultural, Multilingual Contexts

Amanda Datnow, Johns Hopkins University

Our research project, "Scaling Up School Restructuring in Multicultural, Multilingual Contexts" (*Talking Leaves 3*, 1), focuses on the effectiveness of various school restructuring designs in improving achievement of students in schools serving linguistic and ethnic minority students in a multicultural context. We are conducting longitudinal case studies of 13 elementary schools, each of which is implementing one or more nationally regarded school reform models.

Classroom observations are a key component of our study. Through observations, we are able to better understand if and how the reform models are being implemented and adapted by teachers for use with diverse student populations. Observations also help us characterize classroom instruction and organization in the restructuring schools.

For our study, two researchers visit the schools each time, each observing in a different classroom for approximately 90 minutes. We deliberately observe academic subjects. Our observational methodology has four parts.

1) Ethnographic description. Running notes of classroom activities are the first part of our observational strategy. Our emphasis is on documenting the sequence of activities (reform-related and otherwise) in classrooms, paying particular attention to how teachers and students interact and how they derive meaning from classroom interactions. We pay specific attention to how cultural and linguistic identities of participants impact interactions.

2) Reform design implementation. As observers, we assess the reform design's implementation through questions focusing on key classroom elements of the reform design. Using a template approach, we compare the teacher's actual implementation to what the developers intended. These design-specific forms were created with the assistance of descriptive reports and articles about each reform method. The format and length of the forms vary depending on the specificity of the design at the classroom level.

From the Other Side of the Classroom

Ann Gibb, CREDE

Teachers at Starlight Elementary School in Watsonville, CA shared their experiences regarding classroom observations with *Talking Leaves* staff. As a professional development school, Starlight enjoys a partnership with UC Santa Cruz, CREDE, the Santa Cruz New Teacher Project, and the Pajaro Valley Unified School District in the preparation and professional development of undergraduate students, preservice graduate students, interns, paraprofessionals, beginning teachers, and veteran teachers. The schools' focus on equity and academic achievement includes clear principles for educators to become culturally and linguistically responsive. Of Starlight's 760 students, 75% are eligible for free or reduced price lunch, 70% are English language learners, 21% are native English speakers, and 9% are fluent English speakers with a first language other than English. Fourth and fifth grade teachers Shawna Hodovace, Nancy Jackson, Monica Larenas, Kerrin Murphy, Maureen Ryerson, and Jarry Stillman, with principal Noni Mendoz Reis and Professional Development Coordinator/Advisor Irene McGinty, devoted their lunch period to discussing the practice of observation as seen from the other side of the classroom.

These teachers experience classroom observation on several levels: by administrators as an evaluation tool, by their advisor and peers for professional development and support, and occasionally by researchers–some unaffiliated with Starlight–who are gathering data. Most of the discussion focused on the power of observation for professional development, "to help me move my practice forward and to help my students," according to one teacher. The faculty have benefited immensely from a combination of observation, thoughtful dialogue with observers, and self reflection.

At Starlight, the practice of classroom observation for professional development is contextualized to the teacher's and students' needs, and is responsive to the teacher's personal style of learning. For one instructor, classroom observations done at regular intervals by CREDE researcher and professional site coordinator Audrey Sirota were especially helpful because "Audrey has the context. She can connect what I did today with what I did two weeks ago and help me see how I build on a previous lesson, and how I connect the lessons even more." Other teachers request observation and feedback as needed, for help presenting a particular subject, developing a teaching strategy, exploring a new topic, or whenever they need collegial support. In her first year in

3) Effective practice in culturally diverse classrooms. In addition to assessing whether the restructuring model (or some adaptation thereof) is present, we also assess whether the classroom is characterized by effective instructional practices. This is determined through a series of questions guided by theory and research on what constitutes effective classroom practice for culturally and linguistically diverse students. These questions are both a means to discover whether reform features contribute to these effective practices as well as whether teachers incorporate other activities to serve their diverse student populations.

4) The low and high inference observation instrument. We also employ a low and high inference classroom observation instrument called the Classroom Observation Measure (COM). Developed at the University of Memphis and validated in extensive pilot research and other studies of elementary school classroom instruction (Ross et al., 1991), the COM includes both interval (5-minute) coding, holistic ratings, and descriptions reflecting more global impressions of classroom activities. A detailed training manual accompanies the COM. We modified the COM slightly for the purposes of this study.

Our four-prong observational methodology maximizes time spent at each school by allowing us to gather a much wider range of classroom data than if we focused solely on one strategy. By employing multiple kinds of observational data and multiple observers, we are better able to triangulate our data. However, our methodology is time-consuming. In addition, because of our different training and life experiences, we filter what we see through individual lenses, and thus observe life in classrooms differently. Despite these limitations, we believe our approach provides a thorough description of the classrooms we observe and uncovers how classroom participants make sense of the reform models, given their varied social contexts.

Reference
Ross, S.M., Smith, L.J., Lohr, L.L., McNelis, M.J., Nunnery, J., & Rich, L. (1991). *Final report: 1991 classroom observation study*. Memphis, TN: Memphis State University, Center for Research in Educational Policy.

Related Publication
Stringfield, S., Datnow, A., & Ross, S. M. (1998). *Scaling up school restructuring in multicultural, multilingual contexts: Early observations from Sunland County*. Research Report No. 2. Santa Cruz, CA and Washington, DC: Center for Research on Education, Diversity & Excellence. Also available online: www.cal.org/crede/pubs/.

Looking for Findings in All the Right Places

Gil Garcia, Office of Educational Research and Improvement, U.S. Department of Education

Classroom observation seems a simple process. Researchers, parents, administrators, and/or colleagues watch, listen, and make notes (perhaps even record audio or videotape) and reflect on what they observe. But CREDE's conference on classroom observation methods revealed that the process is extremely complex.

Observation strategies used in education research span a wide range, from very quick and superficial to very detailed and long term. However, all methods share two features: the strategies are used to observe an event in its context, and the methods are purposive. As demonstrated through the methods presented in CREDE's conference, purposes can vary: collecting data on pedagogy, providing feedback to teachers, learning about teacher-student interactions, understanding peer dynamics, recording student behavior, or evaluating effects of particular programs.

CREDE's conference was a unique opportunity for researchers to share strategies they have developed or adapted, to learn each other's methods, and to discuss what works, what does not work, and why. The conference created a forum for discussions on multi-faceted observation strategies that increase probabilities for data collection strong in key aspects of school/classroom events and their context dynamics, regardless of the scope of the research. As you will see in this issue, although classroom observation is a complex process, when strategies are well-designed and implemented, results yield abundant and useful data with potential for findings that make a difference in the education of children.

HOT OFF THE PRESS

Implementing Two-Way Immersion Programs in Secondary Schools

This report, by C. Montone and M. Loeb, offers practical advice for two-way immersion programs considering expansion into middle school and high school. It includes an overview of likely challenges, options for meeting these challenges based on the experiences of programs that have already implemented secondary TWI programs, and detailed portraits of seven articulated programs. (EPR 5, $5.00) For ordering information, see www.cal.org/crede/pubs or contact crede@cal.org or 202-362-0700, ext. 258.

The Classroom Observation Schedule and the Teacher Roles Observation Schedule

Hersh Waxman, University of Houston

In the CREDE project, "Improving Classroom Instruction and Student Learning for Resilient and Non-Resilient English Language Learners" (*Talking Leaves 3, 1*), researchers are trying to determine why some Latino English language learners are more successful than others, despite coming from the same backgrounds. In this study, the researchers used two tools to collect data: the Classroom Observation Schedule (COS) and the Teacher Roles Observation Schedule (TROS). This article describes each tool and its advantages and limitations.

The COS is used to collect data on individual students and their behaviors in ongoing classroom instructional-learning processes. The COS is systematic, and designed for documenting several variables: a) interactions with teachers and/or peers and the purpose of those interactions, b) settings in which observed behaviors occur, c) types of material they are working with, and d) specific types of activities in which they engage.

The TROS is designed to describe the nature and patterns of teachers' instructional behaviors. Teachers are observed with reference to: a) the extent of interactions with others, b) the setting in which observed behaviors occur, c) language used, d) interaction purposes, and e) the nature of interactions. The instrument's recording sheet includes these five categories and provides space for recording 10 intervals of these behaviors.

Usually, the TROS is used in conjunction with the Classroom Observation Schedule (COS) because the two instruments complement each other well. With the COS, approximately six students per class are randomly selected for observation. Depending on the study's focus, student selection can be stratified based on variables such as student gender, ethnicity, or high- or low-achieving status. Each student is observed for ten 30-second intervals during the data collection period. The observer watches the student for a 30-second interval and then records the appropriate categories, indicating predominant type of interaction, selection of activity, setting, manner, and language used. If a student works in two or more settings, the researcher records the setting that predominates or occupies the student for the greatest amount of time. Activity types, such as reading, listening, or working on

The Standards Performance Continuum: Measuring CREDE's Standards for Effective Pedagogy

Ruth Soleste Hilberg, **R. William Doherty**, **Peggy Estrada**, **Roland G. Tharp**, & **Georgia Epaloose**, University of California, Santa Cruz

The Standards Performance Continuum (SPC) is a rubric providing a quantitative basis for measuring how well teachers use CREDE's Five Standards for Effective Pedagogy. We designed the SPC to be a reliable and valid measure of the developmental continuum of teacher implementation of these standards. The continuum ranges from no use of the standards, to limited use, to simultaneous enactment of multiple standards. As a developmental continuum, each higher level of implementation subsumes the skills of the lower levels. There are five levels of the SPC, with ratings ranging from 0 to 4:

0) Not Present indicates that the standard is not observed, and no aspect, element, or indicator of the standard is present in the observation.

1) Emerging indicates that although the standard is not observed, some element or aspect of the standard is present in the observation.

2) Developing indicates that key elements of the standard are present, although the standard is not fully enacted.

3) Enacting indicates a complete and accurate enactment of the standard.

4) Integrating indicates the skillful integration of multiple standards simultaneously.

The SPC is currently being used for a broad range of professional development, assessment, and research purposes. For example, researchers and administrators use the SPC to assess teaching practice through live observations, observations of videotaped practice, and review of unit plans and artifacts from teacher professional development portfolios. Researchers are examining the relationship between implementation of CREDE's Five Standards and student outcomes. Teachers use the SPC to facilitate self-and peer-assessment and "partner dialogues"– professional conversations intended to improve the effectiveness of instructional activities. Professional developers use the SPC to assess and guide professional development programs. To learn more about the SPC, visit the CREDE web site (http://www.crede.ucsc.edu/Reports/5.6AERAdoc.html).

written assignments, are not mutually exclusive, and the observer should check as many categories as he or she observed.

The TROS similarly records teachers' classroom behavior in 30-second intervals, during which period the observer records as many purposes and natures of interaction as appropriate. For example, if a teacher uses both questioning and modeling during a 30-second interval, then both categories should be recorded. Similarly, the observer then records the appropriate categories indicating the predominant type of interaction, setting, and language that occurred during the 30-second interval. If a teacher works in two or more settings, the setting that predominates or occupies the greatest amount of time should be recorded.

Since the TROS is a time sampling of a teacher's classroom behavior, teachers are observed non-continuously, facilitating a wider sample of observations. The 10 intervals are spaced for about every 5 minutes (one sweep) throughout the 50-60 minute observation period. The observations of individual students from the COS are typically done between sweeps of the classroom teacher.

Data from the COS and TROS can be used for several purposes. One is as process data, looking for factors that directly or indirectly affect student outcomes. Data may also reveal contextual conditions, including teacher attitudes or behaviors, degree of program implementation variables, teacher background characteristics, or school-level characteristics. Finally, data from both tools can be used to examine and evaluate changes in aspects of student behavior and teachers' classroom instruction.

The COS and TROS produce reliable and valid data. They only require a short training period (usually one half day) and provide excellent feedback to classroom teachers (Waxman, 1995; Waxman & Huang, 1999). The COS and TROS allow for coding of language use (e.g., Spanish or English), an important consideration for studies conducted in linguistically diverse school settings. An advantage of the COS is that it focuses on individual students, so it is less threatening for classroom teachers and enables researchers to investigate important subgroup differences. The large number of variables is a benefit of the TROS.

However, both tools have their limitations. The COS and TROS require the observer to be very focused, using a timer or watch with a second hand. The tools can also be obtrusive. The COS also includes a limited range of variables/activities, sometimes requires parental permission or an attendance sheet from the teacher to observe particular classifications of students, and sometimes requires a larger sample of students. The TROS is more limiting–in 5 classroom visits, researchers can collect data on over 50 students, but only on 5 teachers. The TROS also usually requires a more-experienced educator/observer who understands classrooms and two or three different observation periods to obtain an adequate sample of a teacher's instructional style. 🍁

References

Waxman, H. C., & Huang, S. L. (1999). Classroom observation research and the improvement of teaching. In H. C. Waxman & H. J. Walberg (Eds.), *New directions for teaching practice and research* (pp. 107-129). Berkeley, CA: McCutchan.

Waxman, H. C. (1995). Classroom observations of effective teaching. In A. C. Ornstein (Ed.), *Teaching: Theory into practice* (pp. 76-93). Needham Heights, MA: Allyn & Bacon.

Director*, from page 1*

impatient when they encounter reflex resistance to observation; after all, how can education ever improve if educators do not study classroom processes? And few researchers or teachers have escaped disagreements over the value and meaning of a particular observation system.

One CREDE project is collecting observations from classrooms nationwide to see the relationship between teacher performance of the Standards for Effective Pedagogy and student achievement. The purpose is to guide further research and program development. In other CREDE studies focused on professional development, observers provide direct feedback to teachers, who are themselves concerned with improving their practices, so researchers are well-advised to observe and report on things the teachers themselves are interested in improving.

Some years ago, I led a large educational research-and-development organization, and became interested in studying the development of the organization itself. I provided an anthropologist with access to all our staff meetings, consultations, and supervisory interactions so that we researchers were observed regularly as we worked, just as our teachers were observed in their classrooms. None of us would ever again underestimate the impact of being observed!

I suspect the key to the complexity of classroom observation lies in the value attached to the variable or characteristic being observed. When researchers attach great value to certain observations, and teachers do not, both are likely to fear the influence of the other on later events. Will the principal or the school board or the teacher act in some unforeseen way as a result of this information? On the other hand, when teachers themselves are interested in the events being observed, cooperation flows through a partnered purpose. The work is necessary though complicated in action and in emotion, and it can be useful and gratifying for all. 🍁

The Kentucky Primary Program Component Configuration Map

Ellen McIntyre and **Diane Kyle**, University of Louisville

In 1990, Kentucky's legislature passed the Kentucky Education Reform Act (KERA), requiring comprehensive change to the state's educational system. KERA mandated restructuring kindergarten through third grade classrooms into multi-age, nongraded primary programs by the fall of 1993. These programs were to reflect seven "critical attributes": a) multi-age, multi-ability grouping, b) developmentally appropriate practice, c) continuous progress, d) authentic assessment, e) qualitative reporting, f) professional teamwork, and g) positive parent involvement.

The seven critical attributes are addressed by the Kentucky Primary Program Component Configuration Map, a 4-point rubric developed by Kentucky researchers, including the authors of this article. The Configuration Map allows observers to examine detailed aspects of each critical attribute. For instance, the "developmentally appropriate practice" attribute includes individual rubrics for reading, writing, mathematics, science, social studies, student talk, teacher/student interaction, physical arrangement of the room, and more.

For their study, "Appalachian Children's Academic and Social Development in Nongraded Primary Schools: Model Programs for Children of Poverty" (*Talking Leaves 3*, 1), McIntyre and Kyle selected teachers who were "high implementers" (high scores in most or all of the seven categories). Through their collaborative study, teachers examined the rubric repeatedly, scoring themselves, discussing meanings of the rubric, and comparing their progress to their "scores" on the Standards Performance Continuum and standards for best practice (Zemelman, Daniels, & Hyde, 1998). Conversations around the rubrics seemed to bring teachers to new understandings and new practices. One teacher said: "The configuration map can help us know what areas are 'stretches' for us. The things I think I am a '2' on instead of a '1', I want help in understanding why."

While the instrument has been reliable in evaluating progress in implementating the primary program, to our knowledge no studies show a correlation with increased student achievement, nor have studies compared the instrument to other normed instruments evaluating instruction. For now, the Configuration Map seems best used as a tool for discussion and analysis with teachers. Through reflection and conversation, using descriptions from the rubric as standards for good practice, teachers can increase their understanding about what pedagogical features best meet the needs of their students.

For a copy of the Configuration Map, contact the Kentucky Institute for Education Reform, 101 Taylor Education Building, Lexington, KY 40506. 🍁

Reference

Zemelman, S., Daniels, H., & Hyde, A. (1998). *Best practice: New standards for teaching and learning in America's schools.* Portsmouth, NH: Heinemann.

---

the classroom one teacher struggled "to get what Irene was telling me. Finally I asked her to demonstrate it, and I became the observer. By watching her model, I got what she had been trying to tell me." This cycle of observation, demonstration, collaborative reflection, and repeated observation has become a common form of support.

The feedback process is an equally fluid, open-ended, responsive, and interactive process with the mutual goal of providing explicit support to improve teaching practice, thus supporting students' learning. These faculty enjoy and appreciate being observed because the opportunity to see their teaching from another's vantage point is profoundly useful. All teachers appreciated watching themselves teach on videotape. One teacher used the videotape for building community: "I let my class see it so they can see themselves as part of the class."

These practitioners, and many education researchers, share the same goal–improving teaching practice through classroom observation. Yet, this group of teachers did not readily see the value of being observed by researchers collecting data; a process which was generally viewed as "someone's else's agenda." Observation for the purpose of research was not perceived as a reciprocal process; however, teachers saw ways to make the experience more meaningful for faculty. "I'm an expert in my room," one teacher explained. Researchers could tap into this expertise by collaborating with teachers on research design, talking with faculty about what teachers need and what would be useful for their students before beginning research, conducting research on policy issues relevant to the school, and asking teachers about effective strategies. By combining their experience and skills, and working together on research, teachers and researchers can more quickly achieve their common goal of helping students through improved pedagogy. 🍁

Upcoming Summer Institute

Teaching English Language Learners: Effective Programs and Practices

June 27-29, 2000
University of Connecticut, Storrs Campus

Standards for pedagogy, and effective programs for linguistically and culturally diverse students will be among the topics discussed at a national conference on the education of English language learners at UConn's Storrs campus this summer. The institute has been designed for preservice and in-service teachers at all levels, school and district administrators, and policy makers involved in bilingual and English as a second language education.

The program will include plenary sessions presented by nationally-recognized experts, interactive workshops, and concurrent presentations. Participants will have opportunities to learn, practice, and share effective, research-based strategies. Institute highlights include:

- Choosing and implementing programs that fit school and district needs
- Exploring latest research findings on language learning and academic achievement
- Applying strategies for teaching subjects at various grade levels
- Examining the Five Standards for Effective Pedagogy
- Accessing and using resources of the National Clearinghouse on Bilingual Education

Featured speakers include **Dr. Roland Tharp**, director of the Center for Research on Education, Diversity & Excellence (CREDE); **Donna Christian**, president of the Center for Applied Linguistics (CAL) and expert on two-way bilingual immersion programs; and **Hernan LaFontaine**, former superintendent of schools, Hartford, CT.

The conference is being presented by CREDE, CAL, and the University of Connecticut at Storrs. For registration information, call 860-486-2005 or see http://www.crede.ucsc.edu/Resources/Conferences/suminstitute.html.