

Mapping the Process: An Exemplar of Process and Challenge in Grounded Theory Analysis

by Beth Harry, Keith M. Sturges, and Janette K. Klingner

This article responds to recent calls for greater clarity and transparency regarding methods in qualitative research. On the basis of a 3-year ethnographic study of the overrepresentation of minorities in special education, the authors address key tenets of grounded theory and attempt to reconcile some of the methodological challenges inherent in naturalistic inquiry. They discuss theoretical considerations and use a visual model to illustrate how they applied grounded theory to this complex and sensitive topic. Emphasizing the social nature of decision making in special education, the authors point to the appropriateness of qualitative methods to the investigation of such issues.

The need for public sharing of methodological approaches and concerns has been emphasized by a number of researchers (e.g., Anfara, Brown, & Mangione, 2002; Eisenhart & Towne, 2003) and has been heightened by the current emphasis on replication and generalizability in research. The No Child Left Behind Act (2001) identified replication and generalizability as indicators of “scientifically based research” (SBR), but spokespersons for the National Research Council (Feur, Towne, & Shavelson, 2002) emphasize that a range of methods are required for addressing the complex issues faced by education. In arguing for “understanding and appreciating the multiple perspectives in education in the service of developing a strong, self-regulating culture” (p. 9), these authors concluded that the field should “dispel the myth that science is synonymous with a particular method,” because “the question drives the methods, not the other way around” (p. 8).

Despite these assurances, proposals for qualitative research are experiencing a level of rejection that presages a reversal of the progress made over the past decade in understanding issues that are best studied by such methods. The research that we report in this article is a particularly good example of the type of question that benefits from naturalistic inquiry, and our purpose is both to explicate the method and to show why it is needed in the study of social processes.

The disproportionate placement of minorities in special education has been of concern to the field ever since the issue was first raised by Dunn (1968) and Mercer (1973). Most often, re-

searchers have used statistical methods to examine this topic, and some studies have been very successful in revealing important associations between minority placement and key variables such as socioeconomic status (SES) and size of group in a district (e.g., Oswald, Coutinho, Best, & Singh, 1999). What continues to be missing from this body of research is an understanding of the motives and pressures that drive decision making in this area; in other words, the “how’s” and “why’s” of the process. Without such understanding, practitioners and researchers alike are left to speculate as to the causes of the correlations revealed by statistical studies. Common daily conversations on the topic frequently rely on assumptions about bias or funding patterns, or about minorities, who may be viewed simply as evidencing disabilities at disproportionate rates.

Following Anfara et al. (2002), we focus on the analytic process involved in implementing the key methods and concepts of grounded theory (Glaser & Strauss, 1969; Strauss & Corbin, 1995), while using examples from our findings to illustrate methodological points. We approach our task in the following steps: First, we offer a brief summary of the research project’s purposes, design, and data collection. Second, we present a discussion of some key aspects of grounded theory, explaining how we approached them in our study design. That explanation will be interwoven with a visual model of our analytic process showing the key steps, and a narrative reflecting on the process. Finally, we discuss the key challenges and lessons learned through our efforts.

Project SEARCH: Research Design and Data Collection

Under the auspices of the Office for Special Education Programs (OSEP), we conducted a 3-year study of the referral and placement of Black (various ethnicities, such as African American, Haitian) and Hispanic (various ethnicities, such as Cuban, Nicaraguan) students in special education programs in a large urban school district. The purposes of the project were (a) to investigate whether and, if so, how, the processes used to identify, assess, and place students in high-incidence special education programs contribute to the overrepresentation phenomenon; and (b) to identify referral and placement decision-making processes that successfully mitigate overidentification and overrepresentation while also providing beneficial educational outcomes for students.

The research design was based on an argument proposed by Heller, Holtzman, and Messick (1982), the authors of the first National Academy of Sciences report on ethnic disproportionality in special education. These authors argued that an under-

standing of the issue must be based on a thorough analysis of the processes through which these placements occur, because inappropriate practice at any point casts doubt on the validity of the outcomes. On the basis of this premise, we designed a study that would examine all phases of the referral process in K–3 general education classrooms, from early instruction, through referral, to assessment and placement outcomes. We further conceptualized the issue as one that is influenced by many levels of social ecology, from the most distal to the most proximal, as in the conditional matrix proposed by Strauss and Corbin (1995) and the ecological theory of Bronfenbrenner (1979).

The research was conducted in a funnel-like process over three phases, moving from a description of countywide placement rates and referral/placement policies, to an examination of the implementation of the referral/placement policies in 12 schools, to individual case studies of 12 students. The research team consisted of two co-principal investigators, who were educational researchers trained in qualitative methods, and a senior research associate, trained in applied anthropology. These researchers remained constant throughout the project and were assisted, each year, by a cohort of three or four undergraduate and graduate assistants who were trained by the senior research associate. The team was multi-ethnic and multilingual, including African American, Haitian, Hispanic, and Anglo-American individuals.

Phase 1 (September–April, Year 1)

We collected data on the school district's placement rates and policies by examining relevant statistics and written documents. On the basis of this information, we selected 12 schools purposively, to represent a range of configurations in ethnicity, SES, and rates of referral to special education. Between February and April, we conducted a total of 71 audiotaped interviews with district and school administrators, related service personnel, and teachers. The purpose of these interviews was to gain the views of key players regarding how the placement process worked for minority children and why overrepresentation exists. The interviews were semistructured, in that we had a set of guiding questions, but the style was open-ended and responsive to the lead of the interviewee (Spradley, 1979). The central question, common to all interviews, was simply, "What do you think explains overrepresentation?"

Phase 2 (April, Year 1–June, Year 2)

In addition to continuing interviews (the same type as above) with administrators, teachers, and other support personnel in each school, we conducted observations in all K–3 general education classrooms in each of the 12 schools (with a few exceptions). Using Junker's (1960) outline of participant observation roles ranging from complete observer to complete participant, our early observations leaned more toward the observer-as-participant role, with sufficient participation to assist in the building of rapport. As we moved into intensive observations in target classrooms (described below), we generally became more participatory in order to learn through interaction with students and teachers.

We selected 2 teachers and their students from each school for more extensive data collection, consisting of 8–12 observations in each classroom, along with interviews and informal conversations with the teacher. This process occurred throughout the second semester of the school year. At the end of that phase we conducted

a set of structured "exit interviews" with the 24 target teachers. The method and purpose of these interviews will be reported later.

Phase 3 (September, Year 3–End of Project)

Using the same ethnographic methods as previously described, we conducted in-depth case studies of 12 students, whose cases represented issues arising from the data. We believe that the use of multiple sources and data collection strategies provided us with considerable saturation and triangulation of data. All interview transcripts and summaries, as well as all observational field notes, were entered in the ATLAS.ti qualitative analysis database. Overall, the data collected included

- Transcripts of 272 open-ended individual interviews with students, parents, and school-based and district personnel and field notes on 84 informal conversations.
- Field notes from 627 classroom observations, 42 child-study team meetings, 5 psychological evaluations, 15 special education placement and/or Individualized Education Plan (IEP) meetings, 14 other school-based meetings, and 15 home and community settings relevant to target students.
- Documents such as IEPs; students' work, psychological and other evaluations; school district guidelines and policies; and extant data on special education placement in the school district.

Methodological Considerations in Qualitative Research

In using a visual display, we note that, although such a model is helpful to the researchers themselves and instructive to those desirous of learning the method, any visual representation of a complicated cognitive process is a vast simplification of the way that researchers actually arrive at interpretations. In the case of the code-mapping model that we use in this article, which is similar to that used by Anfara et al. (2002) and which reflects the work of Guba and Lincoln (1985), the main limitation is the tendency for it to appear as a linear process. Our arrangement of the model to be read from bottom to top is an attempt to reflect visually the inductive nature of the process—moving, through several analytic iterations, from the "ground" up. Later in the discussion, we will address the contradictions and ambiguities inherent in the belief that one can be truly inductive.

We acknowledge that no model can represent the intuitive leaps that are an essential part of any analysis. Such intuitions represent moments of insight in which the researcher makes inferences based on what Glaser and Strauss (1969) called "sensitizing" concepts, which allow the researcher to quickly grasp the meaning implicit in social situations. This must be tempered by careful documentation of the researcher's perceptions (Peshkin, 2000) and further tested by a search both for substantiation and for negative cases. As Van Maanen (1988) noted, ethnographic writing has undergone a shift from the "realist" mode to the "confessional" and even to the "impressionist" mode. The first of these reflected the belief that the researcher's job was to capture and report objective reality. However, Geertz's (1983) assertion that ethnographers should focus not only on "native" ways of life but on their own construction of knowledge has been followed by a considerable body of literature challenging the notion of objectivity. Clifford (1988), for example, made a substantial contribution to this literature with a critique of his own ethnographic

writing, which, he argued retrospectively, had feigned objectivity and failed to explain the processes by which accounts and themes were produced. Other examples are accounts by Harry (1995) and Peshkin (1988) that attempt to make transparent the impact of various aspects of researchers' identities on the research act.

Explaining the Grounded Theory Analytic Process: Language and Iterations

Glaser and Strauss's (1967) original conception of grounded theory methodology was framed in terms of a series of iterations that the authors referred to as a process of "constant comparison," in which the researcher moved back and forth among the data and gradually advanced from coding to conceptual categories, and thence to theory development. Strauss and Corbin (1998), in a development of this approach, referred to the first step in the process as "open coding," through which the researcher names events and actions in the data and constantly compares them with one another to decide which belong together. When the actual language of the speaker is used as a code name, the name is called an "in vivo" code; in this article we enclose such codes in quotation marks. In Glaser and Strauss's language, the "basic, defining rule for the constant comparative method" is that, while coding an incident, the researcher should compare it with all previous incidents so coded, a process that "soon starts to generate theoretical properties of the category" (p. 106). For example, in our data, a teacher, distressed about the large number of children in her class, exclaimed, "Oh, no! So many kids!" We compared the properties of the situation to which she was referring with a statement by another teacher: "There are 23 [exceptional education] kids lined up at my door." Noting that both teachers were complaining about the number of children they were expected to teach, we assigned both statements the code Class Size. As the analysis continued, every event or comment similar to these was examined to decide whether it referred to the same complaint.

The next step is to group the discrete codes according to conceptual categories that reflect commonalities among codes. Strauss and Corbin refer to this as "axial coding," reflecting the idea of clustering the open codes around specific "axes" or points of intersection. It is important to note that, when engaging in categorizing/axial coding, these properties are being identified through the interpretive lens of the researcher, who is already beginning to abstract meaning from the data. For example, we decided that the code Class Size fit into a category of codes that we called Classroom/Teacher, which referred to classroom issues that were challenging to the teacher. It is evident that in assigning class size to this category we were going a step beyond the descriptive; we were no longer simply noting that class sizes were perceived to be large but were indicating that large class size was one of the issues that teachers struggled with, as they did with behavior management or children's cultures. However, as we became more aware of the role of administrators in making decisions about class size, this code seemed better related to a category called Administrative. In the earliest iteration of our analysis, the code Class Size was subsumed under the category Classroom/Teacher, but in a refinement of the coding system we later categorized it as an Administrative issue. This is not to say that the code could not have been placed in two categories, but only that, as our interpretation of the data developed, we thought it more

useful to separate teacher skills from classroom issues that were beyond the teacher's control.

Strauss and Corbin (1998) refer to the third analytic level as "selective coding," meaning that at this point the researcher treats the various code clusters in a selective fashion, deciding how they relate to each other and what stories they tell. Thus the analyst "constructs . . . a set of relational statements that can be used to explain, in a general sense, what is going on" (p. 145). We find it more explanatory to call Level 3 the "thematic" level, referring to the underlying message or stories of these categories as "themes." It is in seeking the interrelationships between these themes that the researcher begins to build a theory. Strauss and Corbin state that, as an "overarching theoretical scheme" or "central category" is identified, theory development becomes a recursive "search for consistency and logic" (p. 156) that will integrate both complementary and competing strands of evidence. By examining negative cases, filling in poorly developed categories, and accounting for variation, the researcher refines the theory. In Strauss and Corbin's words, "Integration occurs over time, beginning with the first steps in analysis and often not ending until the final writing" (p. 161). Turning to our mapping model (Figure 1), we will now illustrate our analytic process.

Project SEARCH Data Analysis Map

Levels 1 and 2: Open Codes and Conceptual Categories in Initial Interview Data

After the initial 3 months of intensive collection of field data, we devoted the summer to analysis of the data from the Phase I interviews. Those analyses produced what our model shows as Level 1 and Level 2 analyses. The 71 interview transcripts were read by one researcher who, working at the most concrete level, developed a set of "open" codes that labeled the key points being made by the interviewees.

The bottom row in Figure 1 ("Open Codes") indicates the initial set of 125 open codes that are listed in Table 1 (p. 8). We reviewed these codes, comparing and contrasting them to identify the common features among them, in order to cluster them into "conceptual categories." Level 2 (second row from the bottom in Figure 1) contains those categories. For example, 10 codes (Attitude to Testing, Participation, In Denial, Labels, Assistance, Problems, Family Structure, Ethnicity, Education, and Poverty) were all subsumed under the category Families, indicating aspects of parents and parenting that interviewees believed contributed to or protected students from special education placement. This process resulted in a set of 21 categories, which represented the range of information gathered about school personnel and their views of the main contributors to African American and Hispanic children's placement in special education. Rather than include all of these in the Figure 1 model, we have included only the dominant categories and have listed the entire set of 21 categories and their supporting codes in Table 1.

The codes and categories were then shared with the rest of the research team, who gave feedback toward condensing and rearranging some of the categories and clarifying the properties of each code. For example, there was considerable discussion about overlap among the categories Community, School, and Society/Outer Circle, which resulted in careful attention to discriminating between those factors that were closer to the child and those

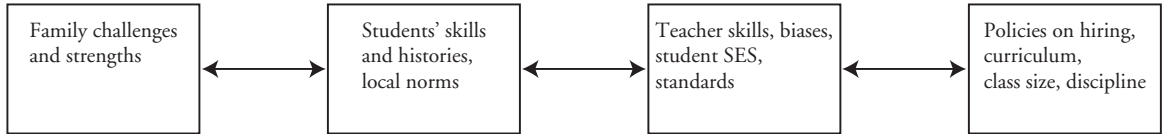
6. Theory:

Influences—A complex set of negative influences contribute to the overrepresentation of minorities in special education. Predominant contributors are the assumption of intrinsic deficit and the requirement for a disability categorization; inequitable opportunity to learn, resulting from poor teacher quality in lower-SES schools and higher standards in higher-SES schools; negative biases against families perceived as dysfunctional; external pressure from high-stakes testing; and subjectivity in referral and assessment practices.

5. Interrelating the explanations:



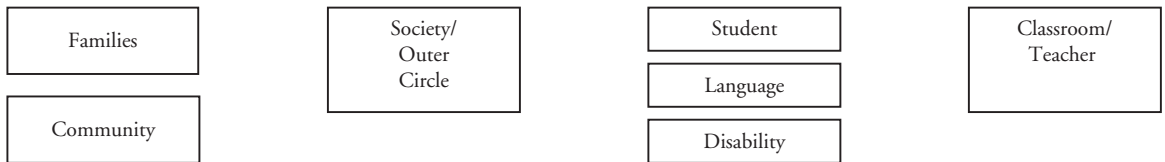
4. Testing the themes (interviews, observations, documents):



3. Themes:



2. Categories:



1. Open Codes: Based on initial interviews.

FIGURE 1. *Data analysis map. The numbers at far left represent the six levels of analysis, moving upward from the bottom of the figure. Two-directional arrows indicate nonlinear connections among items. See Table 2 for more detail on Level 4; see Table 1 for more detail on Levels 1 and 2.*

that were more distal. Thus references to the drug dealers in the neighborhood would be categorized as Community, whereas references to historical patterns of poverty or to federal or state educational decisions would be categorized as Society/Outer Circle.

Testing the codes and categories for clarification and reliability. After reviewing the full set of interview codes and categories, we embarked on a process of testing the codes for clarity and reliability. A sample set of transcripts were read by a team that included one co-principal investigator and 3 graduate assistants, who met twice a week for 3 weeks to compare their separate coding of the same data. This process continued the constant comparison approach, in which a given data point would be compared with another to see if the same code would apply, thus developing consistency in usage of the codes. As a result, some codes were condensed into one and new codes were developed, as commonalities or distinctions among the meanings of similar data points became clearer. For example, under the category, Professional, we initially included two codes: Attitude to Profession and Extra Mile, the former referring to professionals' definitions of the limitations of their roles and the latter referring to specific statements about "going the extra mile," such as buying shoes for

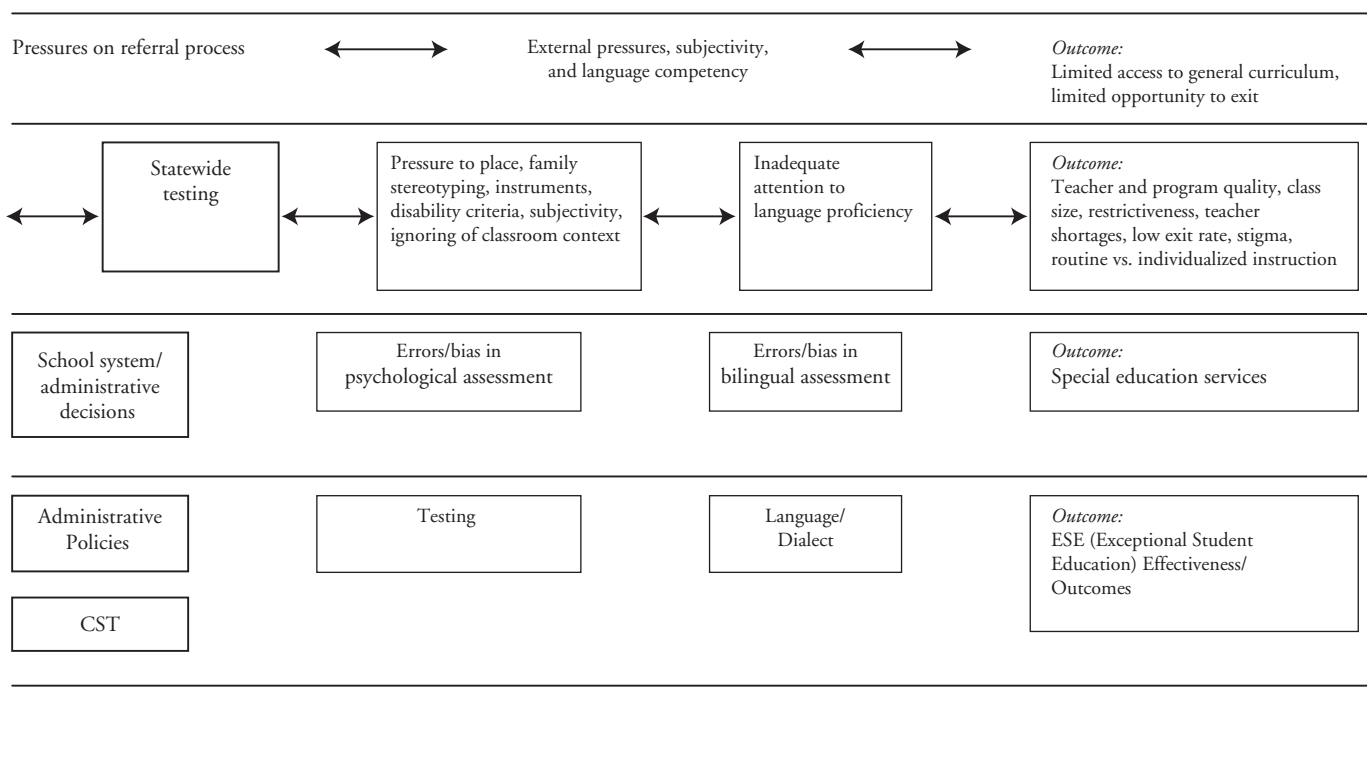
a child. The "in vivo" code Extra Mile, although attractive, proved relatively rare, and it was more economical to subsume it under Attitude to Profession.

We made no attempt to develop a numerical reliability rating, because our goal was consensus, with each point of difference being debated and clarified until the group agreed on appropriate usage of the set of codes.

Level 3: Developing the Themes

The team's next task was to move toward thematic analysis. The essential question at this stage is: "What are the themes embedded in the conceptual categories?" We began by determining which categories were predominant in the data and summarizing their content. Overall, our coding system identified two general types of data: demographic and/or personal information (e.g., school personnel's personal and professional histories, programs offered by the school, and racial or SES demographics of schools and communities), and information about school personnel's perspectives on educational issues, students, colleagues, families, and communities. With the exception of descriptions of Child Study Team (CST) processes, which were detailed at great length, the data on personnel perspectives were the most common, no doubt because our interview approach focused on eliciting the speakers'

Outcomes—The outcome of the process is placement in special education programs of variable quality and, especially in the case of Emotional Disturbance and Educable Mental Retardation, unduly restrictive environments. Such placement is often detrimental because of stigma and limited opportunity to exit. Low-SES, African American populations are particularly vulnerable because of the accumulated effect of biased institutional practices. (Outcomes are represented at far right, below.)



views. This, indeed, was our purpose, and we treated the descriptive data as background information to which we would return later.

Focusing on our participants' perspectives, we found that five categories were predominant in the Phase I interviews: Classroom/Teacher, CST (Child Study Team), Families, Community, and Administrative Policies. We decided also to summarize the data in the Testing and Disability categories, although they appeared only minimally in the data. The reasons for focusing on them were their evident relationship to the CST process and the research team's awareness of the frequency of such explanations in the research literature.

The task of summarizing the content of these categories was distributed across the team and, once the summaries were done, team members reflected on the content and determined what seemed to be the main arguments within each category. We called these arguments "themes," which reflected seven types of explanations for overrepresentation:

1. Family/community influences (including parental participation in children's schooling)
2. External pressures on schools (school district, state, federal)
3. Deficits seen as intrinsic to child
4. Teacher skills/biases
5. School system/administrative decisions
6. Errors/bias in psychological assessment
7. Errors/bias in bilingual assessment

Level 4: Testing the Themes

Level 4 of Figure 1 represents several layers of analysis in which we tested and developed the themes by examining the extent of their relevance to incoming data from interviews, observations, and documents. Level 4 illustrates the analysis with selected subjects commonly found in the data. We remind the reader that the model's visual impression of linearity is not an accurate representation of the complex iterations typical of this level of analysis.

So far, we have argued that our analytic process was "grounded" in the data, because it moved inductively toward hypotheses or explanations of the data instead of the other way around. Nonetheless, we readily acknowledge that researcher interpretation begins early in the process and therefore sheds doubt on the notion that the ultimate theory will be totally "grounded." In educational research, as in many social sciences today, the researcher is often a relative insider in the field, studying a topic that he or she already knows quite well. It would be naïve to think that preconceived beliefs and perspectives will not be brought to bear on the data. However, researcher reflexivity works hand-in-hand with the iterative nature of the research to bring preconceived beliefs into the dialogue, rather than seeking to omit or ignore them. These efforts at transparency in the analysis are supported by prolonged engagement in the field and by testing the emerging interpretation against participants' perspectives, a process sometimes referred to as "member checking" (Lincoln & Guba, 1985).

Table 1
Initial Categories and Codes

<i>Administrative Policies:</i> Concerns Of, Concerns About, Support, Initiatives
<i>Classroom/Teacher:</i> Class Size, Paraprofessionals, Expectations, Rewards, Bias, “Chaos/Mayhem,” Behavior Management, Stress, Planning, Culture, Instruction, Relationships with Students
<i>Community:</i> Demographics, History, Cultures, Community Support, Socioeconomic Status
<i>CST (Child Study Team):</i> Process, Records, Strategies, “Qualify,” Data, Timelines, Race/Ethnicity, Referral Reasons, Members, Paperwork, Retention, Staffing, Student Participation, Effectiveness, Grade at Referral, ED Referral
<i>Disability:</i> “True LD,” “True EMR,” “True ED,” ED–LD Interaction, LD–Language Interaction, “PDK”
<i>ESE (Exceptional Student Education) Effectiveness/Outcomes:</i> Programs, Instruction, Effectiveness, Regular Ed Collaboration, Mainstreaming, Gifted, IEP
<i>Ethnicity:</i> African American, Hispanic, White, Haitian, “Islander”
<i>Families:</i> Attitude to Testing, Participation, In Denial, Labels, Assistance, Problems, Family Structure, Ethnicity , Education, Poverty
<i>Language/Dialect:</i> Limited English Proficient, Dialect, Culture, Race, Interpreter, Deficit
<i>Overrepresentation Theories:</i> Broad Theory (multi-faceted, often lengthy opinions on overrepresentation, from which specific excerpts were coded discretely, using codes in this list)
<i>Professional:</i> Role, Experience, Attitude to Researcher/Research, Attitude to Profession, Development, Stress, Personal
<i>Programs/Academic/Nonacademic:</i> SFA, Summer School, ESOL, Home Language Arts
<i>Research Procedures:</i> Sampling, Observing, Incentives, Research Purpose
<i>Researcher:</i> Level of Comfort, Input, Role, Impressions, Expectations
<i>Resources:</i> Materials, Computers, Textbooks
<i>School:</i> History, Demographics, Scores, Attendance, Funding, Mobility, Safety , Maintenance
<i>Setting and Location:</i> any factual description of physical characteristics
<i>SLP(Speech/Language Pathology):</i> Therapy, Caseload, Race, Types, Pragmatics, Articulation
<i>Society/Outer Circle:</i> Societal Problems (Crime/Poverty), Political Decisions (High-Stakes Testing/Vouchers)
<i>Student:</i> Gender, Readiness, Social/Academic Skill Deficits, Strengths, Culture
<i>Testing:</i> Standardized, Psychological, Adaptive, Bilingual, ED, SLP, FAB, Academic, Medical

Note. Italicized expressions are categories; each category is followed by the code or codes (roman type) that it contains. Expressions enclosed in quotation marks represent the word choice of the interviewee. LD = learning disability; EMR = educable mental retardation; ED = emotional disturbance; “PDK” = “Pretty dumb kids”; IEP = individualized education plan; “Islander” = student from the English-speaking Caribbean islands; SFA = Success For All; ESOL = English for speakers of other languages; FAB = functional assessment of behavior.

Level 4 of the figure indicates our efforts to go a step further than the usual member checking, by using the seven Level 3 themes as a frame for examining subsequent data and by testing that frame against the perspectives of key participants. In other words, we were asking, To what extent do the data support these themes?

As mentioned above, although the theme development focused mainly on the participants’ views that predominated in the data, we decided also to include a topic that was rare in the data—the belief that there may be bias in testing for disability. These data referred to two main aspects of testing—psychological and bilingual. We were struck by the paucity of statements about testing bias in the data, because such explanations are found very commonly in the literature on overrepresentation. The belief that psychological evaluation might be biased against minorities appeared in only 3 of 71 interviews and was offered with considerable caveats. Similarly, the suggestion that English

acquisition difficulties might be mistaken for cognitive impairments appeared 4 times in these data.

In addition to the seven themes, Figure 1 also includes, at the far right of the model, an eighth data set—special education services—which is not an explanation for overrepresentation but a concern as to its outcome. We consider it essential to include these data because poor outcomes of special education placement mean that overrepresentation of minorities is indeed problematic.

Testing the themes. We checked the themes against the data in two ways. At the end of Phase II, after about a semester of extensive observations in 24 target classrooms, we conducted structured exit interviews with the target teachers, using a “card-sort” technique. The themes were printed on separate index cards and were presented individually to each teacher, who then sorted the cards into piles of “agree,” “disagree,” and “irrelevant.” The results of this exercise indicated a clear-cut pattern in which the teachers consistently agreed with the themes that

pointed to parent/community influences and children's intrinsic deficits, and consistently disagreed with those that focused on teacher skills. The responses were split half-and-half regarding administrative and curricular influences.

In the exit interview process we committed one obvious error: So stuck were we on giving priority to the themes that arose from school personnel that we did not include the theme of testing bias. Consequently, the card-sort exercise did not inform us of teachers' views of the role of evaluation and assessment in overrepresentation. Fortunately, the overall data set included numerous other conversations and interviews with the teachers, which produced a largely consistent belief that the assessment process provided an accurate determination of the presence or absence of a disability. Most saw "disabilities" as being either inherent in the child or induced by detrimental family circumstances.

The second way that we tested the themes was to apply them to all interview and observational data. Our questions then became (a) To what extent do we observe evidence of these themes in our data? and (b) What additional themes emerge from our observational data? From here on, both observational and interview data were analyzed using the set of codes and categories representing the entire set of seven themes, while additional codes were developed as needed to capture all that we observed.

This process resulted in a lengthy but much more fine-grained set of codes that detailed various aspects of each theme. We offer selected examples in Table 2. For example, data coded within the Family and Community Circumstances category were analyzed into a subset of nine aspects; data coded within the Teacher Skills category were analyzed into two parts, Instructional Skills and Management Skills, each having its own set of four and five subsets, respectively. In addition to the codes and categories that represented the seven themes, we also applied the "descriptive" codes referred to earlier (e.g., Classroom Structure, Content, and Location, which are not listed in Table 2). It is evident from the examples in Table 2 that the final set of codes, although very detailed, was still rather unwieldy—perhaps one of the reasons that (as we will discuss under limitations) our analysis never attained the level of refinement needed to account for all aspects of the data.

So far, we have explained our mapping of the analytic process as far as Level 4—the development of the themes into a set of explanations for disproportionality—and have illustrated in Table 2 our subsequent reanalysis of the content of those explanations. We found that the data reflected the seven themes very well, with some borne out quite consistently and others containing contradictory perspectives. A central tenet here is this: In applying these seven themes to the interview data, we were essentially analyzing our interviewee's perspectives. In applying the seven themes to the observation data, we were essentially analyzing our own perspectives—asking whether we were seeing evidence of these themes in the events that we observed. Thus we constantly compared our own perspectives with those of our research participants.

In our model, we conceptualize the outcome of the Level 4 analysis as a refinement of the themes produced at Level 3. In the visual display we can list only the kernel of each of the emerging themes. A full explanation, for example, in the case of Family and Community Circumstances, would point to a pattern of negative attitudes toward families, based on perceptions of family structures, lifestyles, and experiences. This pattern included a

Table 2
Refined Coding System (Selected Examples)

<i>Category: Family and Community Circumstances</i>	
	Caregiver Perceptions (of schools and services)
	Caregiver Practices and Knowledge
	Community Differences
	Cultural Knowledge Set
	Economics (micro or macro)
	Family Crisis
	Family Setting
	Family Structure (includes descriptions of family members)
	School Perceptions (of caregivers and community)
<i>Category: Teacher Skills</i>	
	Instructional Skills
	Pedagogical
	Socio-emotional
	Style/Personal Expression
	Technical
	Management Skills
	Behaviorist Techniques
	Group Management
	Socio-emotional
	Style (overlaps with instructional style)
	Unwritten Curriculum
<i>Category: Policy</i>	
	Administrative
	Administrator Perceptions (perceptions of or by administrators)
	Assignment of Students
	Assignment of Teachers
	Class Size
	Full-Service School
	Interruptions (includes students and adults entering, leaving)
	Others in the Room (other than teacher and students)
	Retention Policies
	Scheduling
	Specific Policies
	Teacher Out
	Other
	Curricular
	Computers
	Math
	Reading Programs
	Other Curricular
	Other Required Classes

strong tendency to express more positive attitudes toward immigrant minorities than toward African Americans. Researchers' observational data related to this explanation, however, showed that although all of the families who participated did face serious challenges, they also displayed strengths that served as protective factors for children. (For a full discussion of family issues, see Harry, Klingner, & Hart, in press.)

Level 5: Interrelating the Explanations

At Level 5, with our analysis firmly grounded in extensive, triangulated data, we refer to the themes as "explanations"—emphasizing the power of our analysis to develop a theory that

explains disproportionality. This level consisted of trying to come to conclusions about contradictions within an explanation, such as that noted above regarding perceptions of families, and then comparing across the explanations to see how they related to each other. In the model we call these “interrelated explanations.” For example, in trying to resolve contradictions within the data on family issues, we concluded that the discrepancy between school personnel’s and researchers’ views indicated that school personnel’s beliefs about families were often based on minimal and stereotypical information. Of course, we also considered that researchers’ views may have been conditioned by previous research and biases in favor of families. Nevertheless, in several cases, researchers’ home visits revealed family strengths that directly contradicted the school’s view. For example, a family that had been described as one where the children were “raising themselves” was found to have a clear and effective pattern of older sibling authority and responsibility as well as strong neighborhood support and protection. The latter was evidenced by the fact that when a researcher visited the home in the parents’ absence, a neighbor immediately appeared inquiring as to the identity of the researcher and the purpose of her visit. In another family, a parent whom school personnel described as “retarded” offered the researcher a clear and convincing analysis of the limitation of the services her child was receiving.

In comparing across the Level 5 explanations, our mapping model uses horizontal arrows to indicate that they were all interrelated. Once more, the model is far too condensed to tell the story, and because our purpose here is simply to display and critique the method, we will offer only some key explanations.

Our central finding. One finding is essential to our understanding of the topic: Our interrelation of the explanations showed that no single explanation could stand alone. None could be supported or refuted without reference to another explanation with which it interacted. For research concerned with the nuances of human behavior, we consider this a realistic finding, one that underscores the difficulty of measuring complex social processes. The reader of such research must be willing to trust his or her own evaluation of the data presented, relying on the logic of the report rather than on a simplified formula that assigns specific weighting and specific directionality to each variable.

We elaborate this point with the example of school personnel’s belief that detrimental family circumstance was the main contributor to special education placement. We found that certain indicators of detrimental family circumstances—in particular, knowledge of drug abuse, incarceration, or fragmented family structure—tended to lead school personnel to assumptions that went well beyond the information that they actually possessed. Those assumptions, in turn, led to decisions or interpretations that directly affected children, sometimes in negative ways.

Specifically, we cite the example of one of our case studies: Robert, an African American second-grader who was referred for behavioral difficulties. In this case, school administrators’ knowledge that Robert’s mother had nine children and a history of drug abuse led to an attitude of disdain that was illustrated in many ways. The most negative outcome of that attitude was the administration’s decision to place Robert on half-day placement and require his mother to pick him up at eleven o’clock every morning. One of the key players in the decision stated explicitly, to the

mother and to us, that this action would teach the mother that she had to be responsible for her son’s behavior. This arrangement continued for 5 months, during which Robert’s behavior deteriorated further, until his mother finally objected and Robert was reinstated full-time and sent for psychological evaluation. To the dismay of most members of the evaluation team, the psychologist found Robert to have Attention Deficit Hyperactivity Disorder (ADHD), a condition that made him eligible only for part-time special education services. With an Individualized Education Plan (IEP) that had no requirement for specific behavioral support or intervention, Robert spent several unsuccessful weeks in that setting and was soon referred for reevaluation. Reviewing the case, the same psychologist then determined Robert to be eligible for services as a child with Emotional Disturbance (ED), which resulted in Robert’s removal from the school to a self-contained ED classroom at another school. This was the outcome that most members of the evaluating team had wanted from the beginning of Robert’s case.

We do not argue that Robert had no behavioral problems. Indeed, his mother, whom we observed to be attentive and loving with her children, was very concerned about his behavior and had signed consent for evaluation prior to his being placed on half-day status. We do argue, however, that rather than offering this child appropriate support or behavioral intervention, the school effectively exacerbated his difficulties and increased the likelihood of his ED placement. Thus we have no way of knowing whether more appropriate decision making would have resulted in Robert’s being categorized as ED. We contend also that this situation occurred partly because of school personnel’s attitude toward this mother, whose deferential manner and low social status as a stay-at-home welfare recipient allowed her to be taken advantage of in a way that would not have been possible in other circumstances.

Such findings are not measurable. Yet, in the presence of rigorous data collection and analysis, their trustworthiness (Lincoln & Guba, 1985) is evident. The data presented many complex interactions of this nature. In some the logic of the situation was evident, although complicated. In others the data were so mixed, even contradictory, that the timeframe of the project impeded a more fine-grained analysis.

We find some consolation in Berliner’s (2002) description of the difficulty of disentangling data that are embedded the real-life interactions of schools:

A science that must always be sure that the myriad particulars are well understood is harder to build than a science that can focus on the regularities of nature across contexts. . . . Doing science and implementing scientific findings are so difficult in education because humans in schools are embedded in complex and changing networks of social interaction. (p. 19)

Level 6: Delineating the Theory

Glaser and Strauss (1967) distinguished between formal and substantive theory. The latter term refers to a set of explanations that account for phenomena within a specific or substantive field; in contrast, the term “formal theory” is reserved for theories of processes that can be applied to a broad range of similar topics. In this study, we consider our theory to be at the substantive level, in that the explanations we have derived for overrepresent-

tation refer specifically to the placement of African American and Hispanic students in special education programs at disproportionately high rates in the setting that we studied. To attain the level of “formal theory,” our explanations would need to be tested in various situations to see whether general statements could be made that would hold across similar settings.

We reiterate that the complexity of the findings became the centerpiece of our overall theory: That no single factor, such as teacher bias or intrinsic deficits in children, can account for the phenomenon of overrepresentation. Rather, the problem must be understood as reflecting a complex set of negative influences that include aspects of all seven explanations. Interactions among these factors limit students’ opportunities to learn the academic and behavioral skills required for school success and may increase the risk of special education placement. Unless those barriers are removed, it is impossible to be sure—and it should not be assumed—that special education placement represents true disabilities intrinsic to children. Moreover, the entrenched nature of the barriers (e.g., patterns of hiring that provide the best-qualified teachers to the upper-income schools and the reliance on local or neighborhood funding to support the provision of paraprofessionals to counteract large class sizes) amounts to a pattern of institutional bias that loads the placement process against low-income African American and Hispanic students.

One contrasting school underscores the fact that even the complex portrait of institutional bias does not tell the whole story. In that school, which served a high-income, predominantly White population, African American students were both an ethnic and a low-SES minority. The main challenge they faced was the presence of extremely high academic standards set by the majority group in the school. Our data revealed that African American students were overrepresented in special education placements at this school, at double the proportion of their overall presence in the school population. Yet their academic levels seemed to be higher than those of children referred in the inner-city schools. Thus, despite high-quality instruction, the local norms of the school effectively raised the bar for low-income African American students. Furthermore, our research across the 12 schools indicated that pressures from statewide testing were driving building administrators to remove potential failures from the general education program, because up to that time, the scores of students in special education did not “count” in the school’s ratings. This occurred regardless of the SES level of the student population, because a school rated “D” was as eager to improve its rating as was an “A” school to maintain its rating.

Finally, as the far right of the model shows, our findings regarding the quality of special education services indicated that, in most cases, those programs were overly restrictive and provided students with limited opportunities to return to general education. Several factors contributed to that limitation: (a) unduly large classes that undermined special education’s goal of individualization; (b) teacher shortages and poorly prepared teachers; and (c) in the programs for children with emotional difficulties, a focus on behavior control rather than cognitive development and learning.

Challenges, Dilemmas, Limitations

Although we are satisfied with the general theoretical conclusions of the study, there are several limitations that we must live with,

and from which we have learned lessons that will influence our future work. We mention two here, and we believe that by sharing these challenges we will encourage others to benefit from the lessons we learned.

Our first concern was a challenge that we believe should be treated as a foreground issue in discussions of qualitative methodology, and of grounded theory in particular. The essence of the method is its inductive nature. This inductiveness requires the researcher to approach the data from a perspective of relative neutrality, the main goal being to describe and understand, rather than to evaluate, patterns within and across cultures. However, in its adaptation to the purposes of educational research, quite different goals and assumptions have encroached on the method of grounded theory. Educational research has, for the most part, taken the direction of evaluation, its purpose being to develop state-of-the-art practice. Thus the educational researcher usually approaches the data with a great deal of knowledge about literature on the topic being studied, as well as a set of beliefs regarding what constitutes effective schooling. Furthermore, because most topics in education today have some controversial aspect, the researcher is challenged to pay close attention to his or her personal biases, whether based on ethnicity, culture, gender, SES, or any other aspect of individual identity and experience.

If truly reflective research habits are developed, the apparent dichotomy between neutrality and value-laden perspectives need not be an obstacle in qualitative research. Indeed, we propose that this dichotomy can be re-envisioned as a potential bridge between extant knowledge and the researcher’s grounded insights. In our study, this was evident in the example of the relative absence of the theme of testing bias in our grounded data, which stood in strong contrast to our own hunches and years of debate in the field. Our decision not to include this theme in our card-sort exit interviews reflected our determination to test only the emerging grounded theory. In retrospect, we believe that we should have used this opportunity to obtain a firm check on how the target teachers viewed the theme of assessment, while acknowledging that it had been introduced into the equation more from our own knowledge than from the “grounded” data.

The extent of our concern with the perceived dichotomy between neutrality and preconception was evident in one meeting in which we were discussing whether our selection of cases should be based on tentative hypotheses arising from the data or on issues that have been identified as problematic in the extant research literature. The following paragraph summarizes the dilemma, and the irony in the tone reflects the frustration these conversations sometimes produced:

These questions gave rise to an interesting discussion. [One member] raised the question of whether selection of cases should be driven by a pure grounded theory approach, with issues arising directly from our data, or whether our approach is to be more eclectic, including decisions driven by general and relevant “issues out there” in the field. . . . [Another member] felt that we are committed to some issues a priori, such as to examine race/ethnicity and to examine certain disability categories. [The third member] tried to argue that as long as both of these aspects show up in our data it’s still a “grounded theory” approach. (A central issue in this discussion is the extent to which grounded theory researchers can expect to be “tabula rasa.” . . . To be continued for the next 2 years!)

Our second concern focuses on the indomitable challenge of time in using this methodology. The complexity of the data set made it impossible for us to pursue the numerous possible connections between all the data within the 3-year timeframe of the study.

Our plan to apply the initial seven themes to subsequent data worked well. Conceptually, the idea was sound, and the actual coding of all data, though laborious because of the large data set, proved possible and productive. However, although we met weekly during the entire 3 years of the project, the heavy data collection regime sometimes required us to devote unexpected amounts of time to logistics and trouble shooting, thus detracting from our time for team examination of emerging interpretations and for the design of additional data collection strategies that could have pursued intriguing strands of data. Much of the analysis in Levels 5 and 6 (comparing findings across the 12 schools and delineating the theory) was conducted after the official end of the project, because we were not able to complete data summaries from each of the 12 separate schools until almost the end of the project. Though in its earlier stages our approach to data analysis was generally more formal, such as writing memos toward theme development, by the time we got to the Levels 5 and 6, the time press meant that we were often working in a more informal way, with less formal documentation of the analytic process.

In retrospect, we conclude that we could have achieved a more refined analysis with a sample half this size. We consider this an important lesson learned, and it may be that in struggling against our own years of conditioning to seek larger numbers for greater reliability, we tried to strike too much of a compromise between quality and quantity.

Conclusions

We conclude by returning to the current debate on “scientifically based” or “evidence-based” research. We contend that our data constitute powerful evidence. However, we also believe that there is more art than science in the conduct of research on a topic as complex as the one we tackled. That art includes theoretical sensitivity to emerging data, which guides decisions about further data collection and ways to test emerging theory; interpersonal skill in the creation of a balance between rapport and appropriate distance with research participants; and self-awareness that allows for a critical view of one’s own role in the research. We believe that the contribution made by a study such as this underscores the value of qualitative methods in investigating social processes that are, by their nature, not amenable to enumeration or measurement.

Despite the limitations that we have noted, our findings contribute a new and critically important perspective on ethnic disproportionality in special education: the view that the “multiple determinants” referred to in the correlational studies of Oswald et al. (1999) are not limited to simplistic professional bias, low SES of students, or poor funding patterns, as many suspect. Rather, these determinants include complex interactions that link the most proximal levels of Bronfenbrenner’s (1979) model of ecological systems to the most distal. For example, the implementation of a law such as the No Child Left Behind Act interacts with local norms for children’s achievement, which, in turn,

are affected by local practices such as busing of children from a low-income to a high-income neighborhood. This perspective makes addressing the issue of disproportionality all the more challenging, because it calls for a multilevel and multidirectional attack on the status quo of the educational system. (For specific recommendations, see Harry & Klingner, in press).

In closing, we strongly support the field’s call for more openness in examination of qualitative methods in educational research. Students often feel that they are too much on their own in analyzing their data and that, unlike their peers who engage in quantitative studies, they suffer from the absence of clear-cut formulas. We believe that qualitative study would be limited and weakened by cookbook approaches, but we do agree that open debate and dialogue are essential next steps in the increasing integration of qualitative methods into the repertoire of educational researchers. We believe that the preparation of new members of the academy must include a strong focus on a range of research methods that can address the complex issues created by this nation’s linguistic, ethnic, cultural, and socioeconomic diversity. In the words of the British scientist Sir George Pickering, “Not everything that counts can be counted, and not everything that can be counted counts.”

REFERENCES

- Anfara, V. A., Brown, K. M., & Mangione, T. (2002). Qualitative analysis on stage: Making the research process more public. *Educational Researcher, 31*(7), 28–36.
- Berliner, D. C. (2002). Educational research: The hardest science of all. *Educational Researcher, 31*(8), 18–20.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Clifford, J. (1988). *The predicament of culture: Twentieth-century ethnography, literature, and art*. Cambridge, MA: Harvard University Press.
- Dunn, L. M. (1968). Special education for the mildly retarded: Is much of it justifiable? *Exceptional Children, 35*, 5–22.
- Eisenhart, M., & Towne, L. (2003). Contestation and change in national policy on “scientifically based” education research. *Educational Researcher 32*(7), 31–38.
- Feur, M. J., Towne, L., & Shavelson, R. J. (2002). Scientific culture and educational research. *Educational Researcher, 31*(8), 4–14.
- Geertz, C. (1973). *The interpretation of cultures: Selected essays*. New York: Basic Books.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.
- Harry, B. (1995). These families, those families: The impact of researcher identity on the research act. *Exceptional Children, 62*(3), 292–300.
- Harry, B., & Klingner, J. (in press). *Crossing the border from normalcy to disability: Culturally and linguistically diverse students and the special education placement process*. Teachers College Press.
- Harry, B., Klingner, J., & Hart, J. (in press). African American families under fire: Ethnographic views of family strengths. *Remedial and Special Education*.
- Heller, K. A., Holtzman, W. H., & Messick, S. (Eds.). (1982). *Placing children in special education: A strategy for equity*. Washington, DC: National Academy Press.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Mercer, J. (1973). *Labeling the mentally retarded*. Berkeley: University of California Press.

- Oswald, D. P., Coutinho, M. J., Best, A. M., & Singh, N. (1999). Ethnic representation in special education: The influence of school-related economic and demographic variables. *Journal of Special Education, 32*, 194–206.
- Peshkin, A. (1988). In search of subjectivity—One's own. *Educational Researcher, 17*(7), 17–22.
- Peshkin, A. (2000). The nature of interpretation in qualitative research. *Educational Researcher, 29*(9), 5–9.
- Spradley, J. (1979). *The ethnographic interview*. New York: Holt, Rinehart, & Winston.
- Strauss, A. L., & Corbin, J. (1998). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.
- Van Maanen, J. (1988). *Tales of the field: On writing ethnography*. Chicago: University of Chicago Press.

AUTHORS

BETH HARRY is a Professor at the University of Miami, School of Education, Department of Teaching and Learning, Room 312, P.O. Box 248065, Coral Gables, FL 33124-2040; bebeharry@aol.com. Her

areas of special interest include disproportionate representation of minorities in special education and issues facing the families of culturally/linguistically diverse students.

KEITH M. STURGES is a Research Analyst at the Texas Center for Educational Research, 7703 North Lamar, P.O. Box 679002, Austin, TX 78767-9002; keith.sturges@tasb.org. His areas of special interest include school reform, ethnographic and mixed method research and evaluation, and cultural studies of education.

JANETTE K. KLINGNER is an Associate Professor at the University of Colorado, School of Education, Department of Special Education, 249 Education Building, Boulder, CO 80309-0249; janette.klingner@colorado.edu. Her areas of special interest include disproportionate representation of minorities in special education and reading comprehension strategy instruction for diverse populations.

Manuscript received September 8, 2004

Revision received December 19, 2004

Accepted December 27, 2004

David Thissen, New Editor of *JEBS*

As of January 1, 2005, David Thissen is the editor of the *Journal of Educational and Behavioral Statistics*. Editorial correspondence, including manuscripts, should be addressed to David Thissen, Editor, *Journal of Educational and Behavioral Statistics*, Department of Psychology, University of North Carolina, CB#3270, Davie Hall, Chapel Hill, NC 27599-3270; e-mail dthissen@email.unc.edu.

Book review manuscripts, as well as review copies of new books, should be sent directly to the incoming book review editor, Kimberly Swygert, National Board of Medical Examiners, 3750 Market Street, Philadelphia, PA 19104; e-mail kswygert@nbme.org.

Submissions for the Teacher's Corner should be sent directly to the incoming Teacher's Corner editor, Gregory Camilli, Graduate School of Education, Rutgers University, 10 Seminary Place, New Brunswick, NJ 08903; e-mail camilli@rci.rutgers.edu.