

Part III

Qualitative Data

The Right (Soft) Stuff: Qualitative Methods and the Study of Welfare Reform

Katherine S. Newman

Statistical trends are necessary but not sufficient. To me, statistical trends alone are like a canary in a coal mine—they yield life or death information on the “health” of an environment, but don’t always lead to improvement, causes and corrective actions.

Dennis Lieberman, Director of the Office of Welfare-to-Work
U.S. Department of Labor

In the years to come, researchers and policy makers concerned with the consequences of welfare reform will dwell on studies drawn from administrative records that track the movement of Temporary Assistance for Needy Families (TANF) recipients from public assistance into the labor market and, perhaps, back again. Survey researchers with panel studies will be equally in demand as federal, state, and local officials charged with the responsibility of administering what is left of the welfare system come to grips with the dynamics of their caseloads. This is exactly as it should be, for the “poor support” of the future—whatever its shape may be—can only be fashioned if we can capture the big picture that emerges from the quantitative study of post-Aid to Families with Dependent Children (AFDC) dynamics when many of the nation’s poor women have moved from welfare to work.

This research was supported by generous grants from the Foundation for Child Development, the Ford Foundation, the National Science Foundation, the Russell Sage Foundation, the MacArthur Foundation Network on Socio-Economic Status and Health, and the MacArthur Foundation Network on Inequality and Economic Performance.

Yet as the early returns tell us, the story that emerges from these large-scale studies contains many puzzles. The rolls have dropped precipitously nationwide, but not everywhere (Katz and Carnavale, 1998). TANF recipients often are able to find jobs, but many have trouble keeping them and find themselves back on the rolls in a pattern not unfamiliar to students of the old welfare system. Millions of poor Americans have disappeared from the system altogether: they are not on TANF, but they are not employed. Where in the world are these people? Welfare reform has pushed many women into the low-wage labor market, but we are only starting to understand how this trend has impacted their standard of living or the well-being of their children. Are they better off in terms of material hardship than they were before? Are the benefits of immersion in the world of work for parents—ranging from the psychological satisfaction of joining the American mainstream to the mobility consequences of getting a foot in the door—translating into positive trajectories for their children? Or are kids paying the price for the lift their mothers have experienced because they have been left behind in substandard childcare? And can their mothers stick with the work world if they are worried about what is happening to their kids?

These kinds of questions cannot be resolved through reliance on administrative records. Survey data can help answer some of these questions but without the texture of in-depth or ethnographic data collection. States and localities do not systematically collect data on mothers' social, psychological, or familial well-being. They will not be able to determine what has become of those poor people who have not been able to enroll in the system. They have little sense of how households, as opposed to individuals, reach collective decisions that deputize some members to head into the labor market, others to stay home to watch the kids, and yet others to remain in school. Problems like domestic abuse or low levels of enrollment in children's health insurance programs cannot be easily understood via panel studies that ask respondents to rate their lives on a scale of 1 to 10. Though one might argue that welfare reform was oriented toward "work first" and was not an anti poverty program per se, understanding the nature of material hardship is an important goal for any public official who wants to get to the bottom of the poverty problem. Trawling along the bottom of the wage structure, we are likely to learn a thing or two about recidivism as the burdens of raising children collide with the limitations of the low-wage labor market for addressing the needs of poor families.

If administrative records and panel studies cannot tell us everything we might want to know about the impact of welfare reform, what are the complementary sources of information we might use? I argue in this chapter that qualitative research is an essential part of the tool kit and that, particularly when embedded in a survey-based study, it can illuminate some of the unintended consequences and paradoxes of this historic about-face in American social policy.

From this vantage point, I argue that the “right soft stuff” can go a long way toward helping us to do the following:

- Understand subjective responses, belief systems, expectations, and the relationship between these aspects of world view and labor market behavior;
- Explore “client” understandings of rules, including the partial information they may have received regarding the intentions or execution of new policies;
- Uncover underlying factors that drive response patterns that are overlooked or cannot easily be measured through fixed-choice questionnaires;
- Explore in greater detail the unintended consequences of policy change; and
- Focus special attention on the dynamics shaping the behavior of households or communities that can only be approximated in most survey or administrative record studies that draw their data from individuals. This will be particularly significant in those domains where the interests of some individuals may conflict with others and hard choices have to be made.

The intrinsic value of qualitative research is in its capacity to dig deeper than any survey can go, to excavate the human terrain that lurks behind the numbers. Used properly, qualitative research can pry open that black box and tell us what lies inside. And at the end of the day, when the public and the politicians want to know whether this regime change has been successful, the capacity to illuminate its real consequences—good and bad—with stories that are more than anecdotes, but stand as representatives of patterns we know to be statistically significant, is a powerful means of communicating what the numbers can only suggest.

THE CONTENT OF THE TOOL KIT

A wide variety of methodologies come under the broad heading of qualitative methods, each with its own virtues and liabilities. In this section, I discuss some of the best known approaches and sketch out both what can be learned from each and where the limitations typically lie. I consider sequentially potential or actual studies of welfare reform utilizing:

- open-ended questions embedded in survey instruments
- in-depth interviews with subsamples, of survey respondents
- focus groups
- qualitative, longitudinal studies
- participant observation fieldwork

Where possible, I draw on ongoing research to illustrate the strengths and limits of these methods.

Open-Ended Questions Embedded in Survey Instruments

Obviously the great value of survey research is in its large sample size, its representativeness, and the capacity it provides for statistical analysis and causal inference. Typically the items on survey research instruments are close-ended questions based on fixed-choice response categories or questions that require respondents to rate their reactions on set scales. However, it is not uncommon for survey studies to include a limited number of items that are open ended, where respondents either write short responses in their own words with no guidance from the researcher or speak their minds into tape recorders that generate brief transcripts. Open-ended questions embedded in survey instruments typically follow more cut-and-dried queries (Were you “very happy, moderately happy, moderately unhappy, or very unhappy” with the quality of your child’s care last week?) with “why?” questions designed to learn a bit more about the reasoning behind a respondent’s answer. (What kinds of problems did you encounter with your child care last week?) The value of the follow-up question lies in the ability of the researchers to anticipate all the relevant fixed-choice categories. Where this is particularly vexing, open-ended questions can help to illuminate complex patterns while preserving the strength in numbers that survey research provides. They also sometimes have the secondary benefit of maintaining the engagement of subjects who may otherwise become bored and therefore less attentive to typical survey items.

At least two purposes can be served here. A key advantage to embedding qualitative research inside a survey design is that one benefits from the representativeness and sample size, while preserving the insights afforded by qualitative data. Second, open-ended responses (particularly in pilot studies) can be used to generate more nuanced fixed-choice questions for future surveys. Finally, open-ended responses can be coded and analyzed in much the same way that fixed choice questions are, but now with categories that essentially have been generated by the survey respondents rather than forced on them by the researcher. The new categories are more reflective of the experiences or views of interviewees as they see them. If the subjective understandings of respondents are the issue, this is an appropriate method for capturing them on a large scale.

Embedding open-ended questions has obvious limitations. Because of the expense involved in coding the material, open-ended questions are not always practical in large-scale surveys with thousands of respondents. If cost becomes a significant issue, it may be necessary to code a random subsample of the responses. Questionnaires administered face to face or over the telephone can still utilize open-ended items by having the interviewer record the responses or by

using tape recorders. Problems of thoroughness can be minimized through careful training of interviewers. However, open-ended questions can be problematic in self-administered and mail questionnaires, particularly when one is dealing with respondents who have literacy problems.

Subsample and In-Depth Interviews

When one wants to collect more open-ended data from each subject, it may be appropriate to draw a smaller random subsample of a survey population for longer interviews designed to elicit information on a wide range of topics. A simple random sample or a stratified random sample may be used (assuming the appropriate demographic categories can be identified—for example, groups defined by race, age, family status, or those with children of particular ages) and can be interviewed in situ or in a central location. On the other hand, there may be situations for which it is helpful to select purposeful samples (that may or may not be selected randomly) for in-depth interviews. For example, among those leaving the welfare rolls, we may want to learn more about respondents who have never worked or who have not worked in many years. Pulling a subsample of this kind for an in-depth interview study can yield important insights. Of course, among respondents with literacy issues, using mail questionnaires is problematic anyways.

Studies of either kind can explore in some detail the experience “informants” are having in seeking a job, adjusting to employment, managing children’s needs, coping with new expenses, finding transportation to work, relying on neighbors, and a host of other areas that may shed light on the TANF and post-TANF experience. As long as the subsample is representative, the researcher can extrapolate from it to the experience of the universe in the same way one would generalize from any representative group.

The advantage of the smaller subsample is that it solicits greater depth of knowledge on a larger number of subjects, yielding a more well-rounded perspective than is possible with only one or two open-ended questions. Such a methodology is appropriate when the study aims to understand the intricacies of subjective perspectives or the intertwined nature of family behavior when policy change impacts directly on one household member, but indirectly on other household members. Problems of this complexity can be understood only with a great deal of qualitative information.

The longitudinal study of the Milwaukee New Hope experiment is a good example of the value of this kind of research. New Hope provided low-income families in the experimental group with generous childcare, insurance supports, and earnings supplements to bring them above the poverty line to make it easier to remain in the labor force if they work at least 30 hours a week. Under the direction of Greg Duncan at Northwestern University and Tom Weisner at the

University of California, Los Angeles (UCLA), the New Hope research team developed both a longitudinal panel survey and an embedded ethnographic study¹ that drew mainly on (1) repeated interviews with a representative sample of participants and controls, as well as “outliers” chosen because they appeared to deviate from patterns observable in their data and (2) classical fieldwork (discussed in a later section). From Duncan’s perspective, the blending of “hard” and “soft” data has been critical in understanding program impacts:

New Hope’s qualitative data proved indispensable for understanding the nature and meaning of program impacts. As simple as an experimental design may seem, analyses of experimental impacts are complicated by needs to quantify the key outcomes and isolate program impacts within important sample subgroups. Qualitative data are very helpful in both of these tasks.

One of the most important—and initially puzzling—impacts of the New Hope experiment was on teacher-reported improvements in the behavior of preadolescent boys, but not girls. Boys but not girls in the experimental group were 0.3 to 0.5 standard deviations better behaved and higher achieving than their control-group counterparts. Based on the survey data alone, however, we were unable to account for this gender difference.

Qualitative interviews suggested that interviewed mothers felt that gangs and other neighborhood pressures were much more threatening to their boys than girls. As a result, experimental group mothers channeled more of the program’s resources (e.g. childcare subsidies for extended-day programs) to their boys than girls. Further quantitative analyses of both New Hope and national-sample survey data support this interpretation (Romich, 1999). It is unlikely that this important finding about family strategies in dangerous neighborhoods would have been discovered from the quantitative data alone (Greg Duncan, personal communication, 11/29/99).

The New Hope project also has provided useful analyses that separate the experiences of subgroups of participants who have responded differently to the same program opportunities. Because New Hope mirrors what some of the more generous states have tried to accomplish in their welfare-to-work programs, its experience is useful in parsing the differential impact of these supports for working families. As Duncan suggests in his comments on labor supply and earnings, without the qualitative component, it would have been harder to “unpack” the behavioral differences that distinguish subgroups:

It was clear from the beginning of our quantitative work that program effects on work and earnings were heterogeneous. Roughly one-third of the families at-

¹The design of the qualitative sample in New Hope took a random draw from all program and control cases that fell into the family and child sample (essentially, cases with at least one child aged 0-10 at the point of random assignment). The research team did some stratification before drawing the sample, sorting the list by program vs. control status, then by race. Thereafter, the sampling was random within these cells (see Weisner et al., 1999).

tracted to New Hope were already working more than 30 hours and viewed the program's benefits as a way of making work and family demands more manageable. If anything, experimental/control differences in the labor supply of these families were negative. In contrast, families not working full time at the start viewed New Hope as a way of facilitating a transition to full-time work. On balance, experimental/control impacts on labor supply were positive for these families, although stronger in the first than second year of the program.

Qualitative interviews pointed to important heterogeneity among this latter set of families. Some, perhaps one-fifth, had multiple problems (e.g., drug dependence, children with severe behavior problems, relatives in ill health) that New Hope's package of benefits were not designed to address. Others had no such apparent problems and, in these cases, both experimental and control families could be expected to do well in Milwaukee's job-rich environment.

But a third group, who were only one or two barriers away from making it, profited the most from the New Hope package of benefits (Weisner et al., 1999). Program impacts on the labor supply of families with a small number of barriers were large, and larger in the second than the first year. This key set of findings simply would not have been discovered were it not for the qualitative work (ibid.).

Focus Groups

A popular technique for exploratory research involves the use of focus groups, small gatherings of individuals selected for their demographic characteristics who engage in collective discussion following questions or prompts issued by a researcher acting as a facilitator. Focus groups operate in the native language of the participants and can last as long as 2 hours, providing an in-depth discussion of a topic. They can be used for a variety of purposes. Some researchers rely on focus groups as a means of generating questions they expect to ask in surveys. Others use focus groups as a primary means of data collection. Here the appeal usually lies in the modest expense involved: This is a "quick and dirty" method of gathering data on the subjective responses of program participants.² As a result, focus group studies can often be done on an ad hoc basis if they are not part of an initial evaluation design. A wide range of interested parties—from politicians to business firms—utilize focus groups as a means of "testing the market," particularly where public opinion is at issue.

Of course, the focus group approach has limitations. The contamination of opinion that occurs when individuals are exposed to the views of others can

²When one adds in the costs of transcription, this method may be more expensive than it first appears. However, because it involves a much smaller number of people gathered into one place, the logistics are less burdensome and the sheer amount of data probably more manageable than a large-scale survey.

render the data hard to interpret. When particularly forceful individuals dominate the discussion, the views of more passive participants can be easily squelched or brought into conformity in ways that distort their true reactions. Some people understandably are hesitant to air their opinions on sensitive subjects (e.g. domestic violence, employer misbehavior, criminal behavior) in these types of settings.

Moreover, it is hard to make focus groups representative of a population in any meaningful sense. They must therefore be used purposively or with caution. Focus groups are not a good tool for producing data that will withstand scrutiny for representativeness. What they do provide is a relatively inexpensive and rapid means of learning about underlying attitudes and reactions, an approach that may be informative for officials or scholars looking to design more nuanced research instruments. They are often used as an exploratory tool to help design survey or interview studies because they help to expose important problems that should be subjected to more systematic study. These are important goals for researchers. For program administrators looking for ways to give their staff members insight into the lives of those they may see only in “numerical form,” focus groups can be a means of putting a human face on administrative records.

Some of the limitations of focus groups can be addressed to a modest degree through the careful selection of focus group members. Sensitive subjects may best be addressed by drawing together people who are as similar as possible, who have experienced a common dilemma, in the hopes that the similarities between them will lessen any discomfort. Hence investigators often construct focus groups along the lines of racial or ethnic groups, gender or age groups, or neighborhood groups. The “contamination” of forceful individuals can be limited by the guiding hand of a highly skilled facilitator who makes sure that others have a chance to participate. However, none of these approaches eliminates the difficulties inherent in public discussions of this kind.

Focus groups are therefore probably best used to gather data on community experience with and opinions toward public assistance programs rather than to gather systematic data on individual perspectives. For example, the problems associated with enrollment in children’s health insurance systems probably could be well understood by convening focus groups. Indeed, one of the strengths of the method is that it prompts individuals who may not be able to express themselves easily in a one-on-one setting to recall and describe difficulties they have encountered. Information of this kind is far more textured and complete than fixed-choice questionnaires and can help public officials to address the deficiencies in outreach programs, for example.

Qualitative Longitudinal Studies

Welfare reform is a process unfolding over a number of years, where the before and the after may be widely separated and the “in between” states of at least as much interest as the ultimate outcomes. We have good reason to believe

that families pass through stages of adaptation as their children age, new members arrive, people marry, jobs are won and lost, and the hold of new requirements (work hours, mandated job searches) exert their influences. For this reason, it will be critical that at least some of the nation's implementation research follow individuals and families over a period of years, rather than rest easy with cross-sectional studies. Indeed, one need only look at how the Panel Study of Income Dynamics, the National Longitudinal Study of Youth, or the Survey of Income and Program Participation have altered and enhanced our understanding of income over the lifespan or movements in and out of poverty over time to recognize the value of panel studies of this kind.

These longitudinal studies contain very little qualitative data. The number of sample members and broad coverage of information is expensive so that cost containment often means depth has been sacrificed in favor of coverage. However, anthropologists and sociologists have developed longitudinal interview studies in which the same participants are interviewed in an open-ended fashion at intervals over a long course of time. I have two studies in the field at the moment—one on the long-range careers of workers who entered the labor market in minimum-wage jobs in poor neighborhoods and the other on a sample of working poor families, intended to assess the impact of welfare reform on those who were not the targets of policy change—that utilize this approach. In each case, representative samples of approximately 100 subjects were drawn from larger samples of subjects who completed face-to-face surveys. Thereafter, the smaller subsamples were interviewed at 3-to 4-year intervals, for a total of 6-to 8-years' worth of data collection. Here it has proven possible to capture changes in perceptions of opportunity, detailed accounts of changing household composition, the interaction between children's lives and parents' lives, and the impact of neighborhood change on the fate of individual families. Although the samples are very small by the standards of survey research, the depth and nuance of the data that emerges from such an approach are of great value in opening the "black box" that may resist interpretation in studies based solely on administrative records or fixed-choice instruments.

Qualitative panel studies are, however, labor intensive and expensive for the number of respondents they generate. They ask a great deal from participants who typically have to give up several hours of their time for each wave. Given these high demands, providing honoraria of \$50-100 to ensure participation in interviews is generally important to generate adequate response rates. Such generous honoraria would bankrupt a larger study. Longitudinal interview studies are typically done via the use of tape-recorded interviews, which must be transcribed and possibly translated. Given the nature of the data that studies of this kind are seeking, it is often helpful to employ interviewers who are matched by age, race, gender, and class. This process is not simple. For example, I have developed research teams that were closely matched along race and gender lines, only to discover that vast class differences became quite apparent between respondents

who were poor and living in rundown neighborhoods and students who are clearly middle class in origin and living in far better circumstances.³ Indeed class was often at least as important as race in making a match. The gulf between a professor in her forties and informants in their twenties can be quite substantial just because of the different worlds they inhabit because of their ages.

Not all studies attempt the matching process, and the question of whether it is necessary to find counterparts who are sociologically similar is controversial. For example, Edin and Lein, both white professional women, have done exquisite interview work with women of color on welfare and in low-wage jobs. Other white researchers (myself and some members of my research teams) have had good success despite racial differences. Indeed, it is sometimes easier for informants to reveal sensitive information to outsiders who are perceived as less likely to “spread their personal business” around town (Kathryn Edin, personal communication).

My experience has shown that long-term relationships are easier to develop when racial barriers are minimized and a comfort zone is reached based on perceived similarities.⁴ It is imperative to have staff fluent in the languages of the subjects. It is even more important to invest in training the members of a research team: All the matching in the world will not make up for lack of training, and one should never assume that sharing skin color or gender is sufficient. These requirements add to the costs involved in research.

The quality of the data obtained through well-designed and well-executed qualitative, longitudinal studies can make them well worth the effort. This may be particularly true when one wants to go beyond a scholarly or policy audience to engage either the public or political figures in the exploration of welfare reform. Illustrating statistical trends with “real-life” examples of the dilemmas and success stories of former welfare recipients is of great value in this regard. Researchers should not cede to journalists the entire responsibility for telling the story of welfare reform “with a human face” because reporters rarely select their informants systematically and there is no guarantee that their accounts will be anything more than anecdotal.

Qualitative panel studies can be developed with an original sampling strategy that picks up a representative population based on neighborhood residence or participants and matched controls who participate in a social service program. However, they are probably most valuable when they are embedded in panel

³John L. Jackson's (2001) dissertation, “Doing Harlem,” makes this point very forcefully as he negotiates the vast gulf that separated him as a black man in a doctoral program from his black informants who were far less well educated.

⁴One should not minimize the interference that class differences pose, even when interviewers are matched for race. Graduate students from research universities may have very little in common with welfare recipients, even when they share minority status. The class background gap can loom very large.

studies using a survey design and are therefore subsets of the much larger population of survey respondents that can serve as a better basis for statistical analysis.

This embedding strategy has one disadvantage: If the underlying survey is part of a longitudinal panel study, the selection of a subsample that will be accorded more attention may bias the responses of this group to succeeding waves of the survey. Researchers need to evaluate this possibility, though it need not be a serious flaw. Most surveys seeking to track the consequences of welfare reform are going to focus on “objective” and measurable outcomes: hours worked, income earned, jobs acquired, jobs lost, health insurance enrollment, and so forth. Qualitative studies may yield additional information on how these states of being were reached (job search strategies, barriers to insurance enrollment), but in most instances will not compromise the underlying information in a negative (concealing) direction. The experience of providing more information through open-ended interviews may, in fact, encourage greater revelation among the participants in the qualitative study. Researchers will want to check for any systematic biases that may be emerging and, for some purposes, exclude the subsample from statistical analyses of the survey population.

However, I would argue that the advantages of selecting the qualitative sample from an original panel population far outweigh the disadvantages. “Soft” studies of this kind are often suspect on grounds of representativeness and the value of their contribution dismissed as a result. Although one could, in theory, recruit participants in a qualitative study who are similar to those in the survey population, it is always possible that these “add ons” differ enough from the participants to raise doubt. Hence, in my view, it is a safer bet to draw the qualitative sample from the original research universe and risk the chances that their involvement may adversely alter their responses to a longitudinal study. (Obviously this is not a problem if the underlying survey is cross-sectional.)

Participant Observation Fieldwork

Anthropologists and qualitative sociologists often combine interviews with a large “N” with direct observation of behavior in order to fill in gaps that may emerge using other data collection methods. Fieldwork of this kind frequently involves day-in, day-out contact with a subset of a larger survey population, often resident in the same neighborhood or partaking of a common institutional setting (e.g., a welfare office, a job training program). Informal conversation between researchers and informants or between members of a community (with researchers “on the side”) can be illuminating. Direct observation of behavior is often helpful as a check on what respondents (survey or interview) report about their state of mind or the actions they routinely take, as a fieldworker may see it differently when in situ.

Participant observation data are recorded in the form of daily fieldnotes that must be entered into a database then coded, sorted, and analyzed for patterns of

recurrent behavior or illustrative instances of a pattern that might have popped up in another form (e.g., in a quantitative analysis). This data collection strategy is particularly helpful when researchers are dealing with behavior that might be concealed, easily forgotten, hard to elicit, or simply skewed by the desires informants often have to paint their behavior in a better light.

For some years now, for example, I have been conducting a study of the impact of welfare reform on the working poor in New York City. This is a longitudinal interview study involving 100 families in three ethnic groups across Manhattan, Brooklyn, and the Bronx. Three waves of interviews over a 6-year period provide a detailed sense of the difficulties these families have encountered securing childcare or finding work that dovetails with family responsibilities, even though these families were not the targets of reform *per se* (because they were already in the labor market prior to 1996). The two waves of interviews we have completed thus far indicate that although the availability of jobs has improved and wages are rising, problems remain for the working poor precisely because their wages do not push them above the poverty line. Improvements in their personal circumstances are, in many instances, offset by the extraordinary escalation in costs brought about by the same economic boom that is providing more employment opportunity. Rents are rising everywhere throughout the New York City area, dwarfing the gains these families have made, particularly for those who are not in rent-controlled or Section 8 housing. Family budgets are strained; relatives are doubling up; children are moving back and forth from New York to Puerto Rico because, as parents are preoccupied with work all day, some are having trouble supervising their children. These observations are clear enough in the interviews.

However, these data provide only a sketchy sense of how these dilemmas surface at the neighborhood level and how, in turn, that ecological context impacts the families in our study. Hence we developed a community study component of the project, a year's worth of intensive fieldwork in three New York neighborhoods—one primarily African American, one largely Dominican, and one with a large number of Puerto Ricans as well as other Latino immigrants from Mexico and Central America. For the past 7 months, we have been tagging along beside police officers, sitting in classrooms, visiting with congressmen and church leaders, talking with local employers, and devoting a lot of attention to 12 families drawn from our interview sample who live in these three communities.

Participant observation has been a valuable addition to what we know from the interview data. For example, we have been able to see for ourselves what the teen culture of the communities is like and the ways in which it is influencing the behavior of particular members of the households we study as the parents are occupied at work. We have witnessed the dilemmas of poor working mothers who cannot easily control their sons when they reach adolescence and we know how they adjust their work lives to try to provide more opportunity for surveillance. Having worked with these families over a long period of time—before and

after their reentry into the labor force—we have learned that their capacity to steer their adolescents has declined sharply as the pressures for them to hold jobs increased.

For example, one family we have come to know quite well has a teenage son who is faring poorly in his middle school. When his mother was receiving public assistance, she was able to visit the school during the day to confer with his guidance counselors at length and to learn directly from them (as opposed to the filtered news from the son) that he was in danger of being held back a full year. Now that this mother is working full time, she is unable to exercise this level of involvement. Her family clearly benefits from her earnings: There is less tension over finances in the household and the departure of a paying boarder was less of a cataclysm than it might have been otherwise. However, the mother is worried about what will become of her teenage son in school and now depends on him for information on his progress. He is clearly at risk for dropping out altogether, which may impact his mother's employment stability and will surely influence his own trajectory into adulthood.

Our home-based and daycare based fieldwork also has helped us understand the dilemmas of poor working mothers who have been unable to afford or locate quality child care. The youngest children in some of these households are showing the effects of poor quality care, with some displaying seriously worrying behaviors that their mothers believe are the result of untrained or unconcerned childcare providers (including relatives pressed into service). When we compare these children to their older siblings, most of whom had more attention from their mothers when they were little, the differences are striking. This tells us there is a problem to be understood here, for the good fortune of mothers (most of whom report being happier because they are working) may be paralleled by the declining fortunes of their youngest children, an outcome many of Edin et al.'s (1999) interviewees worried about in advance. If suboptimal-quality childcare remains the lot of the working poor, we may come to understand welfare reform as a story with bifurcated outcomes within the same family: good news for Moms, bad news for kids.

The knowledge we have gained about the work lives of our main informants is complemented by the fieldwork we have done in the neighborhoods. We know, for example, that although opportunities for factory work are very limited in the city itself, that a whole private, off-the-books system of van pools carries Dominican workers out to New Jersey factories where they earn just above the minimum wage. Our observational data have shown that the van pools themselves have become a major source of information on job openings for low-skilled workers. The cost of this reverse commute is fairly onerous for low earners, however, amounting to more than 10 percent of take-home earnings for most users.

The perspectives of service providers, teachers, police officers, local politicians, and employers are equally valuable, for they are in a position to look

beyond the immediate concerns of particular families to assess the consequences of welfare reform for neighborhoods and the institutions within them that must absorb the demands that policy change visits upon them. Service providers, particularly those in the child care and medical care fields, are concerned that they cannot respond adequately to the additional needs that have surfaced since time limits were imposed on federal cash assistance. Medical care personnel in poor neighborhoods continue to report that they have not been able to enroll enough children in Child Health Plus and that they are seeing a steady, and often overwhelming, demand in emergency rooms for treatment of conditions that should have been seen long before they reach this critical point.

Teachers and guidance counselors have noticed that they have a harder time getting parents to pay attention to children's school behavior or academic problems because they are not as available as they once were. The coincidence of welfare reform and the imposition of new state testing standards for children at all levels of the school system has ratcheted up the stakes in classrooms throughout New York City, leaving teachers and school principals even more concerned about bringing those with educational deficits up to speed. Without easy access to parents, this is proving a complex task. Ironically, however, this very demand has spurred the city to provide summer school classes, which have been an answer to many a working mother's prayers for childcare.

Police officers report steep declines in crime and much safer streets in the three neighborhoods we are studying. There is no evidence that this trend is related in any direct way to welfare reform, but it is instead part of a nationwide pattern that experts have yet to understand fully. In New York City, however, the move toward more aggressive policing in minority enclaves has met with mixed responses, as a number of notorious cases involving police violence have shown. On a day-to-day level, however, these pressures have surfaced in a higher level of street surveillance and some resentment of "police harassment" by youth in the families we study who report being told to "move on" when they are talking with friends on the corner. Young men, in particular, feel somewhat less welcome in their own neighborhoods than they once did.

For adults, particularly women and elderly men, these changes have been a blessing. They can pass without as much fear, walk to and from the subway without worrying about being harassed by drug dealers. Some report that the drug trade has moved indoors and off the streets, which makes them feel more vulnerable than before. But, on the whole, they approve of the changes or at least are willing to tolerate the increase in police aggression because it means fewer worries accompanying their ordinary movements.

It remains to be seen whether neighborhood safety will improve to the point where one of the chief worries of women moving off welfare and into work—that their children will not be safe if left unsupervised or will get into trouble in the absence of their mothers—will be assuaged (Edin and Lein, 1997, 1999; Newman,

1999; Anderson, 1999). This reservation has played a key role in the past in keeping mothers out of the labor market. Until now, crime rates seemed responsive mainly to levels of community social capital (Sampson, 1997) that could, in turn, be boosted through the deliberate efforts of stay-at-home mothers and elders. The absence of mothers from neighborhood streets as they head into the workplace renders this strategy less effective. In any case, if crime continues to decline, we may see that a key purpose of welfare reform (to get mothers into jobs) will be furthered by policy changes that had nothing to do with it (through policies such as increased community policing [Winship, in press] or the drive to lower crime rates).

Participant observation in TANF offices and in welfare-to-work programs is an important part of the picture as well. Rank's (1995) study is one of the few that attempted to get inside the culture of the old cash assistance system, and it was very valuable for understanding the perspective of welfare clients as they were processed by caseworkers. We shall have to await a new generation of organizational studies based on similar fieldwork methods in order to understand how the new goals of TANF offices—especially job placement—are being absorbed into a bureaucracy that was designed for entirely different purposes (Ellwood, 1988).

Qualitative research on welfare-to-work programs can tell us a great deal about the job retention problem as well. Watkins (1999) offers a compelling account of the disjuncture that plagues some programs that try to build self-esteem as a means of retaining participants, only to discover that graduates consequently expect much more from the labor market than they actually find. High job turnover rates follow as the frustration of discovering that an "I am somebody" campaign runs headlong into the low-wage labor market where the message may be something closer to "You are not important."

These examples are intended to illustrate the value of contextual information generated through the use of long-term fieldwork. Among other things, this approach provides something close to continuous monitoring of a small sample of families or participants in organizations. Rather than let weeks or years go by between short-term contacts, fieldwork permits ongoing contact and the capacity to check what informants say about their state of mind, their survival strategies, their relations with others, and their neighborhood or institutional conditions against what fieldworkers can observe for themselves and/or learn from "experts" situated in the community.

SAMPLING ISSUES IN QUALITATIVE RESEARCH

The data derived from interview and participant observation projects can be used in at least three ways: (1) to generate hypotheses that might be turned into survey research questions; (2) to complement research based on large-sample statistical analyses; or (3) as an end in and of themselves. These three aims are not

mutually exclusive. The difficulty, of course, with the complementary research and “end in itself” approach is that questions of representativeness are always vexing with very small samples and for most research in this genre, small samples are the only affordable possibility.

My own approach has involved embedding the selection of informants within a larger survey design in order to respond to this concern. In 1995-96, we undertook a survey of 900 middle-aged African Americans, Dominicans and Puerto Ricans in New York City. They were chosen to be representative of ethnically diverse and ethnically segregated neighborhoods, with both high and low levels of household income. From this population, a random subsample of 100 respondents was chosen for in-depth interviews at 3-year intervals (1998 and again in 2001). Finally, 12 individuals—4 from each of the ethnic groups of central concern—living in the three neighborhoods described in the previous section were selected from this qualitative subsample. The choice of these particular 12 people was guided mainly by their employment status and family type, with a mix of single parents and intact couples. This nested design has enabled us to generalize from the families we have come to know best to the population as a whole with which we began.

A similar approach has been pursued by the Manpower Demonstration Research Corporation’s “Urban Change” project, a study of the impact of devolution and the time limits of the TANF system on poor families in four cities: Philadelphia, Cleveland, Miami, and Los Angeles. A multidisciplinary team of social scientists are drawing on “administrative records; cross-sectional surveys of food stamp recipients; census tract-level neighborhood indicators; repeated interviews with Executive Directors of community-based social service organizations; repeated ethnographic interviews with welfare-reliant women in selected neighborhoods; and repeated interviews with and observations of welfare officials and line staff...” (Edin and Lein, 1999:6).⁵

The qualitative interview part of the Urban Change project has been following 80 families from high- and medium-poverty neighborhoods in Cleveland and Philadelphia. Under the direction of Edin at the University of Pennsylvania, this project has thus far collected a large amount of baseline information on a series of topics including:

Aspirations for [women’s lives] and their children; experiences with case workers and the welfare system; knowledge about and attitudes toward welfare reform; income and expenditure patterns; educational and work experiences; family life; attitudes toward marriage and future childbearing; health and caregiving; social support; material hardship; use of social service agencies; and perceptions of the quality of their neighborhoods (Edin and Lein, 1999:6).

⁵See Quint et al. (1999) for more detail on the methodology of the Urban Change project.

Families were chosen for this part of the study by selecting three neighborhoods⁶ in each city with moderate to high concentrations of poverty (more than 30 percent living below the poverty line) and welfare receipt (20 percent or more of families receiving welfare). Ten to 15 families were recruited in each neighborhood by posting notices in the target neighborhoods, knocking on doors, and requesting referrals from community leaders and local institutions. They attempted to guard against the overrepresentation of any given social network by utilizing no more than two recruits through any of these sources. This strategy avoided the liabilities of drawing from lists provided by TANF offices (which would necessarily skew the research toward welfare recipients alone). The strategy also allowed the researchers to present a truly independent face to their informants, untainted by connection to enforcement agencies that could affect their cash benefits.

A strategy of this kind probably overrepresents people who are higher on social capital than some of their more isolated counterparts. They have connections. A strict sampling design from an established list may pick up people who are less "hooked in" to institutional resources or private safety nets and will therefore tell us something about people who confront welfare reform from a socially isolated vantage point as well as those who are more connected. However, the liabilities of this approach are considerable, for it is much harder to disassociate from official agencies when pursuing a sample generated randomly from, for example, a TANF office caseload.

The neighborhood strategy employed by the Urban Change project ensures that the qualitative study includes white, black, and Latino families who are particularly disadvantaged. As Edin and Lein (1999:7) have explained, the design will not pick up welfare recipients who live in mixed-income or more affluent neighborhoods. It is possible that this strategy yields a slightly more pessimistic perspective on the consequences of welfare reform as compared with what we would have seen had the study included the entire range of long term-recipients, many of whom moved off of the rolls with apparent ease as unemployment declined. These are the people whose human capital, including prior work experience, made them relatively easy to place. The Urban Change project will tell us how this transition affected those with less going for them, because their neighborhoods (and the contacts they derive from them) are less likely to provide useful information for job hunting. The communities selected as the focus neighborhoods undoubtedly present safety concerns that mothers will have to consider as they scramble to figure out how to care for their children. In the end, these are the more pressing questions in need of answers, hence the wisdom of the Urban Change project's approach.

⁶Neighborhoods are defined in terms of census tracts. Ranging in size from one to four contiguous tracts, these neighborhoods must meet the poverty, welfare receipt, and racial/ethnic composition (Edin and Lein, 1999:7)

Urban Change is not an ethnographic project in the strict sense of the term. Contact is maintained intermittently with the target families, often utilizing telephone interviews in place of face-to-face contact. Intervals of contact are approximately 6 weeks, though this varies by the informants' situation. Nonetheless, it will provide a very rich database, spanning the before and after of the imposition of time limits, that will tell us an enormous amount about the challenges women and their families have faced in transitioning from public assistance to the world of work. The size and ethnic diversity of the sample (including poor whites, often overlooked in studies of the poor), the multicity approach, and the fusion of administrative records, expert perspectives, and the inclusion of welfare-reliant families in communities with varying levels of poverty will help to address many of the more important theoretical questions before us, especially the consequences of race and ethnic differences, neighborhood effects, and human capital differences in the unfolding of welfare reform.

Angel, Burton, Chase-Landsdale, Cherlin, Moffitt and Wilson are in the midst of a similar study of welfare reform and its consequences, the Three-City Study. This project involves a survey, which began in 1999, of 2,800 households from poor and moderate income. The sample is divided between TANF recipients and those who do not receive these benefits. It is restricted to households with young children (younger than age 4) and those with children between 4 and 14. A developmental study of 800 of these families who have children ages 2-4 will be embedded in this larger design. This embedded study will include interviews with caretakers and the fathers of these children.

The Three-City Study also has an ethnographic component directed by Burton. The study will follow 170 families to track how welfare policies affect the daily lives and neighborhood resources of poor families. In-depth interviews will be conducted over the course of 2 years and will cover topics such as the respondent's life history and daily routines. This component also includes diary studies and observations of the participant when she goes to social service offices for assistance. (Winston et al., 1999). The great advantage of the three-city study is the way in which the ethnographic sample is nested inside a larger, more representative survey sample and contextual data set that can analyze neighborhood variables, state- and local-level employment data, and the repeated interviews and family assessments in the child development portion of the project.

This project has an enormous budget and is therefore the "Cadillac" model that few other studies of welfare reform will be able to match. Nonetheless, it is theoretically possible to use a rich fieldwork approach as long as the resources for this labor-intensive form of data gathering are available. Few social scientists would disagree that moving from macrolevel findings based on surveys to the most microlevel data drawn from fieldwork, with mid-range interviews and focus groups in between, is the best possible approach for preserving representativeness but building in the richness of qualitative research.

Few research projects will be able to match the scale of the Urban Change and the Three-city projects. Indeed, even my own more modest study of 100 families in one city required a substantial research budget and a rotating team of fieldworkers willing to commit a total of more than 6 years to the enterprise. Of course, not all studies of welfare reform need to be as long in duration as the ones described here. For state and local officials whose aim is less to explore the theoretical questions that motivated these studies and more to learn in depth about the family management problems of their caseloads, it may be possible to arrange with local universities to organize neighborhood-based research projects that will provide “snapshot” versions of the same kinds of questions.

Another sampling strategy involves the use of “snowball” samples that attempt to capture respondents who share particular characteristics (e.g., low-wage workers or welfare-reliant household heads) by asking those who meet the eligibility criteria to suggest friends or neighbors who do as well. Some classic studies in the annals of poverty research have used snowball samples to great effect (e.g., Lillian Rubin’s *Worlds of Pain*, Elliot Liebow’s *Tally’s Corner*). More recently, Edin and Lein’s *Making Ends Meet* relies on referrals from a variety of sources, including the personal contacts of individuals already in their study population, to build a sample in four cities. The defining feature of a snowball sample is that it gathers individuals into a sample that have some acquaintance with those who are already involved. Multiple snowball techniques seek to maximize the heterogeneity of the sample, while single snowballs maximize the homogeneity of the sample. Neither approach results in a sample that is genuinely random, though the former seeks diversity while the latter explicitly seeks purposive groups.

Snowballs can be bound tightly to a particular network, as was the case in *Tally’s Corner*, or can guard against the possibility that membership will not represent truly independent cases. When the object of study is densely connected webs of friends and relatives, it is important to capture naturally occurring social networks. In this case, the initial selection of the key informant needs to pay attention to representativeness. Thereafter, however, there will be nothing random about the study participants: They will be selected members of the original informant’s trusted associates.

For example, in my recent study of the working poor in central Harlem (Newman, 1999), a representative sample of workers in fast food restaurants formed the core of the research, but a selected subsample was central to a final phase of intensive participant observation that focused on the survival strategies of 10 households and the social networks attached to them. The ten key informants were selected to represent the racial and gender diversity of the universe of workers. Branching out from there, in concentric circles around the 10 key informants, we took in the friends, neighbors, schoolmates, teachers, preachers, distant relatives, and street contacts of these individuals. Hence, although the original subsample was representative, the snowballs grew around them because the

purpose of the study was to learn about how these households managed the many challenges of low-wage work in naturally occurring contexts (school, home, church, extended family, etc.). Ultimately, perhaps as many as 500 additional people were included in this phase of the research, though they were hardly a random sample.

Others have used snowballs to generate the “master sample.” However in this situation it is important to guard against the possibility that network membership is biasing the independence of each case. Some snowball samples are assembled by using no more than one or two referrals from any given source, for example. Edin and Lein’s (1997), *Making Ends Meet* is a good example of a partial snowball strategy that has made independence of cases a high priority. Initially, they turned to neighborhood block groups, housing authority residents’ councils, churches, community organizations and local charities to find mothers who were welfare reliant or working in the low-wage labor markets in Boston, Chicago, Charleston, and San Antonio. Concerned that they might miss people who were disconnected from organizations like those who served as their initial sources, Edin and Lein turned to their informants and tried to diversify:

To guard against interviewing only those mothers who were well connected to community leaders, organizations and charities, we asked the mothers we interviewed to refer us to one or two friends whom they thought we would not be able to contact through other channels. In this way, we were able to get less-connected mothers. All in all we were able to tap into over fifty independent networks in each of the four cities (1997:12).

Using this approach, Edin and Lein put together a heterogeneous set of prospective respondents who were highly cooperative. Given how difficult it can be to persuade poor people who are often suspicious of researchers’ motives (all the more so if they are perceived as working for enforcement agencies), working through social networks often can be the only way to gain access to a sample at all. Edin and Lein report a 90 percent response rate using this kind of snowball technique. Because this rate is higher than one usually expects, there may be less independence among the cases than would be ideal under random sample conditions, but this approach is far preferable to one that is more random but with very low response rates.

Sample retention is important for all panel studies, perhaps even more so for qualitative studies that begin with modest numbers. Experience suggests that studies that couple intensive interviews with participant observation tend to have the greatest success with retention because the ethnographers are “on the scene,” and therefore have greater credibility in the neighborhoods from which the interview samples may be drawn. Their frequent presence encourages a sense of affiliation and participatory spirit into studies that otherwise might become a burden. However, my experience has shown that honoraria make a huge difference in sample retention when the subjects are poor families. I have typically

offered honoraria of \$25–\$100, depending on the amount of time these interviewers require. Amounts of this kind would be prohibitive for studies involving thousands of respondents, but have proven manageable in studies of 100, tracked over time. The honoraria demonstrate respect for the time respondents give to the study.

Though design features make a difference, retention is a problem in all studies that focus on the poor, particularly those that aim at poor youth. The age range 16–25 is particularly complex because residential patterns are often unstable and connections between young adults and their parents often fray or become less intense. Maintaining contact with parents, guardians, or older relatives in any study dealing with poor youth is important because these are the people who are most likely to “stay put” and who have the best chance of remaining effective intermediaries with the targets of these longitudinal studies. Retention problems are exacerbated in all studies of the poor because of geographic mobility. One can expect to lose a good 25–40 percent of the respondents in studies that extend over a 5-year period. This may compromise the validity of the results, though it has been my experience that the losses are across the board where measurable characteristics are concerned. Hence one can make a reasonable claim to continued representativeness. Such claims will be disputed by those who think unmeasured characteristics are important and that a response rate of 60–75 percent is too low to use.

Coding Issues

Qualitative research of any kind—open-ended questions embedded in surveys, ethnographic interviews, long-term fieldwork with families or “neighborhood experts”—generates large volumes of text. Text files may derive from recorded interviews, which then must be transcribed verbatim (a costly and time-consuming proposition), or from field notes that represent the observer’s account of events, conversations, or settings within which interactions of interest routinely occur. Either way, this material is generally voluminous and must be categorized to document patterns of note.

Anthropologists and qualitative sociologists accustomed to working with these kinds of data have developed various means for boiling them down in ways that make them amenable to analysis. At the simplest level, this can mean developing coding schemes that transform words into numeric representations that can be analyzed statistically, as one would do with any kind of close-ended survey data. Turning to the Urban Change project, for example, we find that initial baseline open-ended interviews show that respondents are hoping that going to work will enable them to provide a variety of opportunities for their children. Mothers also report that they expect their social status to rise as they depart welfare and note that their children have faced taunting because of their participa-

tion in AFDC; they trust the taunting will cease once they are independent of state support. These findings come from tape recorded interviews intended to capture their prospective feelings about moving into the labor market some 2 years before the imposition of time limits. These responses can be coded into descriptive categories that reflect the variety of expectations respondents have for the future, or the hopes they have expressed about how working will improve their lives.

Most qualitative interview instruments pose open-ended questions in a pre-defined order. They also may allow interviewers some latitude to permit informants to move the discussion into topic areas not envisioned originally. Within limits, this is not only acceptable, but it is desirable, for understanding the subjective perspectives of the respondents is the whole aim of this kind of research and the instrument may not effectively capture all the relevant points. However, to the extent that the original format is followed, the coding can proceed by returning to the responses that are contained in approximately the same "location" in each interview transcript. Hence, every participant in our study of the working poor under welfare reform was asked to talk about how their neighborhood has changed in the past 5 years. Their responses can be categorized according to the topics they generally raised: crime declining, gentrification reflected in rising rents, new immigrant groups arriving, and so forth. We develop codings that reflect these routine responses in order to be able to draw conclusions such as "50 percent believe that crime has declined precipitously in their neighborhood" or "20 percent object to police harassment of their teenage children."

However, we also want to preserve the nuances of their comments in the form of text blocks that are "dumped" into subject files that might be labeled "attitudes toward the police" or "comments on neighborhood safety." Researchers then can open these subject files and explore the patterned variety of perspectives on law enforcement or the ways in which increasing community safety have affected the patterns of movement out of the home or the hours that mothers feel comfortable commuting to work. When qualitative researchers report results, we typically draw on these blocks of text to illustrate the patterns we have discovered in the data, both to explore the nuances and to give the reader a greater feeling for the meaning of these changes for the informants. To have this material ready at hand, one need only use one of a variety of text-processing programs, including Atlas.ti, Nud.ist, and Ethnograph, each of which has its virtues.⁷ Some proceed by using key words to search and then classify the text. Others permit the researcher to designate conceptual categories and then "block" the text with boundary markers on either side of a section so that the entire passage is preserved. It is even possible to use the indexing capacities of standard word-processing pro-

⁷For helpful reviews of these software packages, see Barry (1998) or "QDA Overview" on the web at http://www.quarc.de/body_overview.html.

grams, such as Microsoft Word 6.0 and above, which can “mark” the text and dump it into subject files for later retrieval.

Most qualitative projects require the analyst both to digest the interviews (which may be as long as 70 pages or more) into subject headings and to preserve the flow of a single informant’s interview through summaries that are preserved by person rather than by topic. I typically maintain both kinds of qualitative databases, with person-based summaries that condense a 70-page text to 5–6 pages, offering a thumbnail sketch of each interview. This approach is of primary value to an academic researcher, but it may not be as important to practitioners who may be less interested in life histories for their own sake and more concerned with responses to welfare reform *per se*.

Practical Realities

Qualitative research is essential if we are to understand the real consequences of welfare reform. It is, however, a complex undertaking, one not responsive to the most pressing information needs of local TANF officials for whom documenting the dynamics of caseloads or the operation of programs in order to improve service is so critical. Yet the information gleaned from qualitative research may become critical to understanding caseloads or program efficiency, particularly if rolls continue to fall, leaving only the most disadvantaged to address. If the pressure to find solutions for this harder-to-serve population grows, it may become critical for administrators and policy makers to figure out new strategies for addressing their needs. This will not be easy to do if all we know about these people is that they have not found work or have problems with substance abuse or childcare. We may need to know more about how their households function, about where the gaps are in their childcare, about the successes or difficulties they have experienced in accessing drug treatment, or about the concerns they have regarding the safety of older children left unsupervised in neighborhoods with crime problems.

Is this information challenge one that federal and state officials should move to meet? Will they be able to use this information, above and beyond the more normative studies they conduct or commission on caseloads in their jurisdictions? To answer this question, I turn to several interviews with officials at the federal and state levels whom I’ve asked to comment on the utility of qualitative data in their domains. Their observations suggest that the range of methods described in this paper do indeed have a place in their world and that the investment required to have this material “at the ready” has paid off for them in the past. However, the timing of these studies has everything to do with the resources available for research and the information demands to which officials have to respond. For some, the time is right now. For others, qualitative work will have to wait until the “big picture” based on administrative records and surveys is complete.

Dennis Lieberman, Director of the Department of Labor's Office of Welfare to Work, is responsible for demonstrating to Congress and therefore to the public at large that the programs under his jurisdiction are making a significant difference. As is true for many public officials, Lieberman's task is one part politics and one part policy science: political in that he has to communicate the value of the work this program accomplishes in the midst of competing priorities, and scientific in that the outcomes that show accountability are largely "bottom line," quantitative measures. Yet, as he explains below, this is a complex task that cannot always be addressed simply by turning to survey or administrative records data:

One of the major responsibilities I have is to demonstrate to the Congress and the American people that an investment of \$3 billion (the size of the welfare to work grants program) is paying off. Numbers simply do not tell the story in its entirety or properly. Often times there are technical, law-driven reasons why a program may be expanding or enrolling slowly. These need to be fixed, most often through further legislative action by Congress.

From a surface perspective a program may appear as a poor investment. Looking behind the numbers can illuminate correctable reasons and present success stories and practices whose promise may lie buried in a statistical trend. As an example: one of the welfare to work program criteria (dictated by statute) would not allow service providers to help those individuals who had a high school diploma. We were able to get that changed using specific stories of individuals who were socially promoted, had a high school diploma (but couldn't read it), and were in very great need. Despite all this, they were walled out of a program designed specifically for them. A high school diploma simply did not lift them out of the most in need category. The numbers showed only low enrollment, appearing at first glance like recruitment wasn't being conducted vigorously enough (Lieberman, 1999).

As this comment suggests, qualitative work is particularly useful for explaining anomalies in quantitative data that, left unsolved, may threaten the reputation of a program that officials have reason to believe is working well, but that may not be showing itself to best advantage in the standard databases.

These evaluations are always taking place in the context of debates over expenditures and those debates often are quite public. Whenever the press and the public are involved, Lieberman notes, qualitative data can be particularly helpful because they can be more readily understood and absorbed by nonspecialists:

Dealing with the media is another occasion where numbers are not enough (although sought first). Being able to explain the depth of an issue with case histories, models, and simple, common-sense descriptions is often very helpful in helping the press get the facts of a program situation correct. There is a degree of "spin distrust" from the media, but the simpler and more basic the better. This, of course, also impacts on what Congress will say and do.

However, as Tom Moss, Deputy Commissioner of Human Services for the State of Minnesota, points out, the very nature of political debate surrounding welfare reform may raise suspicions regarding the objectivity of qualitative work or the degree to which the findings it contributes should be factored into the design of public policy:

Many legislators would strenuously argue that we should not use public resources for this kind of exhaustive understanding of any citizen group, much less welfare recipients. They would be suspicious that perfect understanding is meant to lead to perfect acceptance—that this information would be used to argue against any sanctions or consequences for clients.

I would argue that qualitative data is no more subject to this objection than any other research method and that most officials recognize the value of understanding the behavior of citizen groups for designing more effective policies. Whether officials subsequently (or antecedently) decide to employ incentives or sanctions is generally guided by a theory of implementation, a view of what works. The subsequent research tells us whether it has worked or it hasn't, something that most administrators want to know regardless of the politics that lead to one policy design over another. If incentives produce bad outcomes, qualitative work will help us understand why. If sanctions backfire, leading to welfare recidivism, for example, even the most proreform constituencies will want to know how that comes about. Unintended consequences are hard to avoid in any reform.

For this reason, at least some federal officials have found qualitative data useful in the context of program design and “tinkering” to get the guidelines right. Focus groups and case studies help policy makers understand what has gone wrong, what might make a difference, and how to both conceptualize and then “pitch” a new idea after listening to participants explain the difficulties they have encountered. Lieberman continues:

I personally have found qualitative data (aside from numbers) as the most useful information for designing technical assistance to help grantees overcome program design problems, to fix processes and procedures that “are broken,” to help them enrich something with which they have been only moderately successful, and to try something new, which they have never done before.

My office often convenes groups of similar-focus programs for idea sharing and then simply listens as practitioners outline their successes, failures, needs, and partnerships. We convene programs serving noncustodial fathers, substance abusers, employers and others. We have gotten some of the most important information (leading to necessary changes in regulation or law) this way.

Gloria Nagle, Director of Evaluation for the Office of Transitional Assistance in the State of Massachusetts, faces a different set of demands and therefore sees a slightly different place for qualitative work. She notes (personal communication, 11/30/99) that her organization must be careful to conduct research that is

rigorous, with high response rates and large representative samples in order to be sure that the work is understood to be independent and scientific. Moreover, because collecting hard data on welfare reform is a high priority, her office has devoted itself primarily to the use of survey data and to the task of developing databases that will link various administrative records together for ongoing tracking purposes. However, she notes that the survey work the organization is doing is quite expensive (even if it is cost effective on a per-case basis) and that at some point in the future the funds that support it will dry up. At that point, she suggests, qualitative data of a limited scope will become important:

Administrative data are like scattered dots. It can be very hard to tie the data together in a meaningful way. Quarterly Unemployment Insurance (UI) earnings data and information on food stamps might not give a good picture of how people are coping. For example, what about former welfare recipients who are not working and not receiving food stamps? How are they surviving? We can't tell from these data how they are managing. When we no longer can turn to survey data to fill in the gap, it would be very useful to be able to do selective interviews and focus groups.

Nagle sees other functions for qualitative research in that it can inform the direction of larger evaluations in an efficient and cost-effective fashion:

Qualitative research can also be helpful in setting the focus of future evaluation projects. In this era of massive change, there are many areas that we would like to examine more closely. Focus groups can help us establish priorities.

Finally, she notes that focus groups and participant observation research is a useful source of data for management and program design purposes:

I can also see us using qualitative research to better understand internal operations within the Department. For example, how well is a particular policy/program understood at the local level? With focus groups and field interviews we can get initial feedback quickly.

Joel Kvamme, Evaluation Coordinator for the Minnesota Family Investment Program, is responsible for the evaluation of welfare reform for the state's Department of Human Services. He and his colleagues developed a collaboration with the University of Minnesota's Center for Urban and Regional Affairs; together these groups designed a longitudinal study of cases converted from AFDC and new cases entering the state's welfare reform program. Kvamme found that resource constraints prevented a full-scale investment in a qualitative subsample study, but the groups did develop open-ended questions inside the survey that were then used to generate more nuanced close-ended items for future surveys in the ongoing longitudinal project. He notes the value of this approach:

For the past 15 years, Minnesota really has invested in a lot of research and strategic analysis about what we should be doing to help families.... Yet, it is our most knowledgeable people who recognize that there is much that we do

not know and that we may not even know all the right questions. For example, we have much to learn about the individual and family dynamics involved in leaving welfare and the realities of life in the first year or so following a welfare exit. Consequently, in our survey work we are wary of relying exclusively on fixed-choice questions and recognize the usefulness of selective open-ended constructions.

Resource constraints alone were not the sole reason that this compromise was adopted. As Kvamme's colleague, Scott Chazdon (Senior Research Analyst on the Minnesota Family Investment Program Longitudinal Study), notes, the credibility of the research itself would be at stake if it privileged open-ended research over the hard numbers.

It is a huge deal for a government agency to strive for open-endedness in social research. This isn't the way things have historically been done.... We were concerned that the findings of any qualitative analyses may not appear "scientific" enough to be palatable. State agencies face somewhat of a legitimacy crisis before the legislature and I think that is behind the hesitance to rely on qualitative methods.

Between the reservations the research team had about qualitative work and the recognition they shared that close-ended surveys were not enough, was a compromise that others should bear in mind, as Chazdon explained:

We ended up with an extensive survey with quite a few open-ended questions and many "other" options in questions with specific answer categories. These "other" categories added substantial richness to the study and have made it easier for us to write answer codes in subsequent surveys.

"Other" options permit respondents to reject the close-ended categories in favor of a personally meaningful response. The Minnesota Family Investment Program (MFIP) Longitudinal Study made use of the patterns within the "other" responses to design questions for future close-ended studies that were more likely to capture the experiences of their subjects.

A more comprehensive opinion poll of federal and state officials on the program and on the research evaluation side would no doubt generate other perspectives. Suffice to say for the moment, there is potential for qualitative data to take their place in the arsenal of research approaches needed in order to understand what welfare reform has really meant over the long haul.

CONCLUSION: FORMING RESEARCH PARTNERSHIPS

Given the complexities of this style of research, it probably would be most effective for state agencies to provide requests for payments to which local universities can respond as part of their public service and training activities (as Minnesota already has). Students are a good source of research labor and often are very interested in the problems of the poor. Sociologists, demographers,

political scientists, and anthropologists all can be drafted to assist state officials in understanding how welfare reform is unfolding. With proper planning, long-term panel studies that embed qualitative samples inside a large-scale survey design can be conducted in ways that will yield valuable information to policy makers and administrators. Whether these embedded subsamples are representative of the whole survey universe or purposive samples designed to understand one particular category (e.g., welfare leavers, single mothers with young children), these projects can be of great value. Utilizing this kind of partnership has the advantage of independence from the enforcement agencies with whom TANF participants may be reluctant to cooperate. Because most states have a network of public universities distributed throughout the territory, one can use their location to generate appropriately diverse research populations—urban/suburban/rural, multiple ethnic groups, neighborhoods with different levels of poverty, and areas with higher and lower levels of unemployment, could be among those most important to represent.

Research units of state agencies can also invest in in-house capacities for qualitative research. Even when research resources are tight, making sure that ethnographers and interviewers are part of the team is an important management decision. This may appear to be a “frill,” but it actually may save the day when survey results cannot explain the findings on recidivism or childcare. The presence of ethnographers and interviewers in federal agencies is commonplace now. For example, the Census Bureau maintains a staff of anthropologists and linguists who study household organization in order to frame better census questions. In past years, the Bureau has employed teams of ethnographers to conduct multicity studies of homeless populations to check underrepresentation in the census. As devolution progresses, it will be important to replicate this expertise at the state level in the field of welfare reform.

Whether research partnerships or in-house teams are chosen, the greatest success undoubtedly will be achieved when qualitative research is embedded inside quantitative studies that are either cross-sectional or longitudinal panel studies. The fusion of the two approaches provides greater confidence in the representative nature of qualitative samples, and the capacity to move back and forth between statistical analyses and patterns in life histories renders either approach the richer for its partner.

REFERENCES

- Anderson, Elijah
1999 *Code of the Street*. New York: W.W. Norton.
- Barry, Christine A.
1998 Choosing qualitative data Analysis Software: Atlas/ti and Nudist compared. Sociological Research Online, 3. <http://www.socresonline.org/uk/socresonline/3/3/4.html>
- Edin, Kathryn, and L. Lein
1997 *Making Ends Meet: How Single Mothers Survive Welfare and Low-Wage Work*. New York: Russell Sage Foundation.

- 1999 *My Children Come First: Welfare-Reliant Women's Post-TANF Views of Work-Family Tradeoffs, Neighborhoods and Marriage*. Paper presented at the Northwestern University/University of Chicago Joint Center for Poverty Research Conference, "For Better or Worse: State Welfare Reform and the Well-Being of Low-Income Families and Children, Washington, DC, September 16-17.
- Ellwood, David
1988 *Poor Support*. New York: Basic Books.
- Jackson, John L.
2001 *Harlemworld: Doing Race and Class in Contemporary America*. Chicago University of Chicago Press.
- Katz, Bruce, and Kate Carnavale
1998 *The State of Welfare Caseloads in America's Cities*. Washington, DC: Brookings Institutions Center on Urban and Metropolitan Policy.
- Newman, Katherine
1999 *No Shame in My Game: The Working Poor in the Inner City*. New York: Knopf/Russell Sage.
- Quint, J., K. Edin, M.L. Buck, B. Fink, Y.C. Padilla, O. Simmons-Hewitt, and M.E. Valmont
1999 *Big Cities and Welfare Reform: Early Implementation and Ethnographic Findings for the Project on Devolution and Urban Change*. New York: Manpower Demonstration Research Corporation.
- Rank, Mark
1995 *Living on the Edge: The Realities of Welfare in America*. New York: Columbia University Press.
- Romich, Jennifer L.
1999 To Sons and Daughters: Bargaining, Child's Gender, and Resource Allocation in Low-Income Families. Presented at the annual meeting of the Midwest Economics Association, Nashville, TN.
- Sampson, Robert, S., Raudenbush, and F. Earls
1997 Neighborhoods and Violent crime: A multilevel study of collective efficacy. *Science*, 277:918-924.
- Watkins, Celeste
1999 Operationalizing the Welfare to Work Agenda: An Analysis of the Development and Execution of a Job Readiness Training Program. Unpublished manuscript. Department of Sociology, Harvard University.
- Weisner, T.S., L. Bernheimer, C. Gibson, E. Howard, K. Magnuson, J. Romich, and E. Leiber
1999 From the Living Rooms and Daily Routines of the Economically Poor: An Ethnographic Study of the New Hope Effects on Families and Children. Presented at the biannual meetings of Society for Research in Child Development. Albuquerque, April.
- Weisner, Thomas, et al.
1999 Getting closer to understanding the lives of economically poor families: Ethnographic and survey studies of the New Hope experiment. Poverty Research News. The Newsletter of the Northwestern University/University of Chicago Joint Center for Poverty Research. Based on the Joint Center's Workshop on Qualitative and Quantitative Methods, Chicago, June.
- Winship, Christopher, and J. Berrien
in Should we have faith in the churches? Ten Point Coalition's effects on Boston's youth violence. In *Managing Youth Violence*, Gary Katzmann, ed. Washington, DC: The Brookings Institution Press.
- Winston, Pamela, Ronald J. Angel, Linda M. Burton, P. Lindsay Chase-Lansdale, Andrew J. Cherlin, Robert A. Moffitt, and William Julius Wilson
1999 *Welfare, Children, and Families: A Three City Study*. Baltimore, MD: Johns Hopkins University Press.