



**Models of Integration: The Promises and Pitfalls of Combining
Quantitative and Qualitative Research**

Andrew S. London*
Syracuse University

Saul Schwartz
Carleton University

Ellen K. Scott
University of Oregon

March 30, 2004

An earlier version of this paper was presented on the panel on integrating qualitative and quantitative research methods at the 24th annual APPAM Research Conference, Dallas, Texas, November 7-9, 2002.

*Please direct all correspondence to: Andrew S. London, Associate Professor of Sociology, Center for Policy Research, Syracuse University, 426 Eggers Hall, Syracuse, NY, 13244-1020; (315) 443-5067; aslondon@maxwell.syr.edu. This paper was written with support from Manpower Demonstration Research Corporation's Next Generation Project, which is funded by the David and Lucile Packard Foundation, the William T. Grant Foundation, and the MacArthur Foundation.

Introduction

There is substantial interest in the promises of mixed methods research for improving our understanding of policy impacts. For the past several years, researchers working on MDRC's Project on Devolution and Urban Change and the Next Generation Project have been attempting to use quantitative data (survey and experimental) in combination with qualitative data derived from lengthy, open-ended interviews to understand how welfare reforms affect low-income women, their children, and their families.

In this paper, we begin by reviewing some of the key contributions of these efforts. We first discuss several examples of how combining survey and ethnographic data have raised questions about the prevalence estimates derived from the survey and the limitations of static, point-in-time estimates for understanding the efforts low-income women make to maintain "food security" or mitigate material hardships in their lives, or the ways that social desirability biases may affect survey response.

In the second section of this paper, we discuss how ethnographic data has been combined with analyses of data emerging from experimental evaluations of welfare-to-work programs.¹ We draw on two examples that illustrate how the triangulation of evidence can help expand our thinking about how policies targeted at adults might be affecting children and adolescents, and why those effects might be occurring.

In the final section of this paper, we discuss how our thinking has evolved over the course of our work, and some of our current (and still evolving thoughts) about the promises and pitfalls of combining qualitative and quantitative research methods. Like

¹ We plan to include in subsequent iterations of this paper a discussion of a third example of the integration of qualitative data with experimental data. This example will be drawn from the Canadian Self-Sufficiency Project (Zabel, Schwartz, and Donald 2002).

Gibson and Duncan (2000), we believe there are potential synergies to be gained from the careful combination of mixed research methods, and we identify several factors that we think will enhance the value of combining qualitative and quantitative methods in program evaluations. However, we also believe that all methodologies have their limitations, and the goals of quantitative and qualitative researchers sometimes legitimately differ. We believe we must be careful to specify the types of questions that these different types of research methodologies can and cannot answer, and the kinds of contributions they can and cannot make. To fail to do so risks raising expectations too high.

Data Sources

The analyses discussed in this paper draw on findings from various studies involving survey (cross-sectional and longitudinal), ethnographic (cross-sectional and longitudinal), and experimental evidence. We briefly describe each of the studies and their data sources below.

Many of the analyses discussed below draw on data from the Project on Devolution and Urban Change (see Quint et al. 1999 for details additional details about the Urban Change study). The Urban Change study is being conducted under the auspices of Manpower Demonstration Research Corporation. It is a multi-component study of the impact of welfare reform on poor families with children and the neighborhoods in which they live. The study is being conducted in four large urban counties: Cuyahoga, Ohio (Cleveland); Los Angeles, California; Miami-Dade, Florida, and Philadelphia, Pennsylvania. Among the various components of the study are a longitudinal survey and a longitudinal ethnographic study. Some of the studies we

discuss below (Polit, London, and Martinez 2000; 2001; Brock, Coulton, London, Polit, Richburg-Hayes, Scott, and Verma 2002) draw on quantitative estimates derived from the Urban Change survey (sometimes from all four sites and sometimes from Cleveland only). All of the ethnographic data come from the ethnographic studies. One study draws on baseline data from all four sites (Polit, London, and Martinez 2001), while another draws on baseline data from Cleveland and Philadelphia only (Polit, London, and Martinez 2000). One study draws on baseline and first follow-up data from Cleveland and Philadelphia (Morris, Scott, and London *forthcoming*). One study draws on three rounds of data from Cleveland (Scott, London, and Myers 2002), and another draws on all four rounds of data from Cleveland (Brock et al. 2002).

As background, the first wave of the Urban Change survey involved in-person interviews in 1998-1999 with women who in May 1995 were single mothers receiving welfare and/or food stamp benefits and who were living in neighborhoods (census tracts) characterized by high rates of poverty (30 percent or more of households) or welfare receipt (20 percent or more of households). Using data from administrative records, 1,000 women per site who met these criteria were randomly selected. In-person interviews were completed in 1998-1999 with 3,960 women (78.6 percent of those sampled three years earlier). Follow-up interviews with the same women were completed in 2001. Only one report that is discussed below (Brock et al. 2002) draws on data from the 2001 interview, and that is for Cleveland only.

The ethnographic study involves 3-4 semi-structured, in-person, interviews over a three-year period with approximately 40 women per site. These women lived in neighborhoods of concentrated poverty similar to the neighborhoods in which the survey

respondents lived; however, no woman participated in both the survey and ethnography (i.e., the samples were not linked). Racial/ethnic criteria were also used to select respondents for the ethnographic studies; in Cleveland and Philadelphia, the cities represented in the reports discussed below, all of the women in two neighborhoods per city were African American, while in one neighborhood per city, they were white.

The structured survey and open-ended ethnographic interviews covered similar topics. However, the ethnographic interviews yield richer, narrative data about how families are coping and what they are experiencing. The longitudinal survey and ethnographic data from Urban Change cover similar time periods, which enhances the extent to which they are complementary.

Under the auspices of Manpower Demonstration Research Corporation's Next Generation Project, researchers have also attempted to integrate ethnographic data from the Cleveland and Philadelphia sites of the Urban Change Project with experimental findings emerging from ongoing analyses of data from eight experimental evaluations of welfare-to-work programs. These eight experiments are described in detail in Morris, Huston, Duncan, Crosby, and Bos (2001) and Gennetian, Duncan, Knox, Vargas, Clark-Kauffman, and London (2002), but will not be described in detail here.

Ethnographic Data Suggests the Under-Estimation of Material Hardships, Domestic Violence, and Health Problems in Survey Data

As noted above, having complementary survey and ethnographic data has proved useful for scratching beneath the surface of prevalence estimates and for thinking about the dynamic processes that are both revealed and masked by static, point-in-time survey

estimates. Below, we discuss four examples that illustrate this outcome of having both qualitative and quantitative data available for analysis.

Under-Estimating Food Insecurity

Polit, London, and Martinez (2000) used data from the 1998 Urban Change survey and the baseline ethnography from Cleveland and Philadelphia to examine food security and hunger. Food security in the survey was measured with the U.S. Department of Agriculture's Household Food Security Scale (HFSS). The HFSS consists of 18 items that can be used to classify respondents into one of four categories: food secure, food insecure without hunger, food insecure with moderate hunger, and food insecure with severe hunger (Carlson, Andrews, and Bickel 1999). The scale has been judged to be reliable and valid for both individual and population uses (Frongillo 1999), and has been administered by the U.S. Bureau of the Census's Current Population Survey since 1995. The women in the ethnographic sample were not administered the HSFF; however, they were queried about: food expenditures; use of Food Stamps, WIC, and other publicly-funded food programs; food deprivations; and the use of emergency food services. These data were used to classify women into one of the food security categories with good inter-coder reliability. About half of the women in the survey and ethnographic samples respectively were classified as food secure.

Although the survey and ethnographic samples are not linked and the measurement strategies differed, having both data sources available provided important insights that allowed these researchers to raise questions about the meaning of food security as measured by the HSFF. Polit, London, and Martinez (2000) note that the term food secure might in some cases be misleading because most of the mothers in the

ethnographic sample who were classified as food secure expended considerable energy and pieced together numerous strategies to make sure that there was adequate food for themselves and their children. Women who were food secure talked a lot about: managing their paperwork with the welfare office to maintain their Food Stamps; careful shopping that sometimes involved the purchase of day-old breads or older, reduced-price meats; and routine reliance on friends or relatives for meals or loans that enabled them to feed their families. Arguably, reliance on these strategies are less necessary or extensive among non-poor families who are food secure, suggesting that the label “food secure” masks considerable variation in the efforts families expend to achieve food security and in the precariousness of their ability to maintain this status. Thus, it is possible that the HSFF misclassifies some poor women as food secure when in fact the array of coping strategies they use reflect a daunting struggle that most families who are food secure never face.

The ethnographic data also revealed other issues of relevance to the interpretation of the HFSS categories. In keeping with the conceptual definition of food security as having adequate access to food “without resorting to emergency food supplies” (Anderson 1990), all of the women in the ethnography who used food banks were classified as food insecure. However, it is possible that these women were in fact food secure, partially as a result of their strategic and planned use of the food bank resources available in their communities. This suggests that there might be some misclassification in the HSFF resulting from assumptions about what the use of food banks means and the extent to which poor women are active agents in strategically constructing food security for their families.

Polit, London, and Martinez (2000) note that there were hints in the ethnographic data that women were proud of their ability to feed their children. They might not always be able to pay their bills, but they made sure there was adequate food on the table. There were also hints in the data that some women feared that if their children were not adequately fed, the children would be taken away from them. If there is pride associated with maintaining adequate food, and shame or fear in not being able to do so, it is possible that these concerns would lead some women to avoid endorsing survey items that are indicative of food insecurity. Such social desirability biases would contribute to the under-estimation of food insecurity as measured by the HFSS.

Under-Estimating Material Hardships More Generally

The themes articulated above were further investigated in Cleveland using the 1998 and 2001 survey data and longitudinal ethnographic data. In their analysis of changes over time in material hardships in Cleveland, Brock, Coulton, London, Polit, Richburg-Hayes, Scott, and Verma (2002) reported that from 1998 to 2001 there were significant improvements in several housing and health care hardships among women in the survey sample, and that for only one indicator (unmet need for dental care) did hardship worsen for the sample as a whole. Levels of food hardships in the survey sample were high and unchanged in 1998 and 2001.

Although there were improvements over time in families' economic circumstances, and many women in the survey continued to be reliant upon safety net services at the second interview, material hardships were common and well-above national norms in 2001. However, the ethnographic data suggested that that the prevalence of hardships might have been even higher in the absence of these women's

considerable efforts to cope with, prevent, and reduce the hardships they and their families face. This was the case even among some of the women who substantially increased their incomes when they left welfare for work. For example, Linda, a 33-year-old, white, married mother of three left welfare in 1999 and began working full-time in a job that initially paid an annual salary of \$21,000. She remained in that job for the rest of the study but reported that she still used a food bank on occasion. Similarly, Debbie, a married African American woman with three children, who in the second year of the study entered a job that paid \$8.75 per hour, said that she too was struggling. She said that she had received a shut-off notice for one of her utilities but that the utility was not, in fact, turned off. Nevertheless, Debbie expressed concern about how she was going to pay her upcoming bills --- and indicated how she had managed to keep her utilities on: “They give you up to three months [to pay the overdue balance]; like the second month they give you a shut-off notice, then you get an extra month to try to get it paid, or pay something on it. And now they’re sending me the final bills from where I used to stay, plus the bills I have accumulated here, and it’s like, ‘how am I gonna pay these’? So, I’ll probably just send them what I can.”

Besides going to food banks, women in the ethnographic sample reported numerous other strategies for preventing or reducing food hardships for their families. These included but were not limited to maintaining their connection to the Food Stamp program; cutting back their food intake so their children could have more food; smoking to reduce appetite; eating reduced quality and less expensive foods; receiving food from or eating with family or friends; careful shopping and use of coupons; and stealing. In addition to making partial payments to prevent a utility cutoff, women reported numerous

other strategies to cope with housing problems. These include applying for Section 8; living with relatives (either in the same housing unit or in a housing unit owned by a family member); fighting with landlords to get things fixed; fixing things themselves when they could afford to do so; living with problems unfixed, often for years; borrowing money from friends or family; juggling bills; applying for housing subsidies and emergency assistance; or doing without utilities or phones for periods of time. At times women moved --- either within the same neighborhood or to another neighborhood --- to try to achieve better housing or neighborhood conditions. Sometimes this strategy was successful, but sometimes it was not. Women's strategies for dealing with unmet needs for medical and dental care included: self care; paying out-of-pocket; going to the emergency room; seeking care but not paying the bills; trying to ignore the problem and not seek treatment for it; borrowing medications from others; and stealing medications (by opening packages of over-the-counter medications in pharmacies and taking either the entire contents or a few pills).

In summary, there was evidence in the survey of favorable changes over time across multiple indicators of material hardship. However, the ethnography suggested that some caution was needed when interpreting the point-in-time prevalence rates. Point-in-time estimates mask the substantial efforts that women are making over time to prevent hardships in their families. Thus, even though absolute levels of hardship remained high according to the survey data, it is likely that the survey data under-state the extent to which material deprivation impinges upon the lives of poor women and their children.

Under-Estimating Physical Health Problems

A third example of potential under-estimation was reported by Polit, London, and Martinez (2001; Chapter 4). These investigators used data from the Urban Change survey and ethnography to examine women's physical health status. In the survey, the Short Form 12 Health Survey (SF-12) was used to provide a generic, multidimensional measure of health status (Ware, Kosinski, and Keller 1996). The SF-12 is a widely-used, 12-item scale that includes six physical health and six mental health items. It has been shown to have adequate reliability and validity in various populations and age groups (Ware, Kosinski, and Keller 1996).

One of the six physical health items included in the SF-12 asks respondents to rank their health as excellent, very good, good, fair, or poor. Self-reported or subjective health status, measured in this way, is considered to be one of the best indicators of health status; it correlates with both acute and chronic diseases, and it predicts health services utilization and mortality (Miilunpalo et al. 1997; Mossey and Shapiro 1982; Ross and Mirowsky 1995). In the ethnography, women were asked an open-ended question about how their health has been in general in the past year. This question was followed by more open-ended probing about health problems and concerns.

Approximately 26 percent of the survey respondents indicated that their health was fair or poor, the majority of the rest indicated that their health was good or very good, and 18.6% rated their health as excellent. By comparison, only 7.7 percent of U.S. women aged 18-44 years rate their health as fair or poor, and 35.1 percent indicated that their health was excellent.

Although the women in the Urban Change survey sample were considerably less healthy than the national population of similarly-aged women, as indicated by this self

report, there is some indication from the ethnography that these estimates may be conservative. Just as most women in the Urban Change survey indicated that their health as good to excellent, when they were asked how their health was in general, most women in the ethnography said it was “good.” However, several respondents who initially said they were in good health went on to describe relatively serious health problems once the interviewer probed further. For example, one woman who said her health was “pretty good” went on to describe a serious foot injury that would require surgery, a back injury, asthma, a recent really bad case of bronchitis, and recent surgery to have a cyst removed from her breast. In recounting these health problems, the woman repeatedly downplayed their significance. When she told the interviewer her back injury was not better, she said: “but I just keep going, I don’t let it get me down.” When talking about her asthma, she said: “So, I just keep going, I mean, I try not to think about it.” It was clear from the interview that this woman was trying to minimize the severity of her health problems. This woman was caring for a child with serious mental health problems, and worked at least in part to get some time away from him. Admitting the severity of her health problems, with the potential implication that she would have to withdraw from the work she valued highly for a variety of reasons, led her to under-report her health problems. Thus, the ethnographic data suggest that some women may be reluctant to reveal health problems, which may result in the under-estimation of serious health problems in this population.

Under-Estimating Ongoing Domestic Violence

The fourth example of potential under-estimation can be pieced together from two different research reports on domestic violence. The first reported baseline survey and

baseline ethnographic findings from the four Urban Change sites (Polit, London, and Martinez 2001). The ethnographic data were used to illustrate the types of abuse women faced, but not to raise questions about the survey estimates. The second reported longitudinal ethnographic data from the Cleveland, Ohio site of Urban Change (Scott, London, and Myers 2002).

It is well-documented that poor women are at disproportionate risk for domestic violence, and that many women on welfare have a history of abuse (Tolman and Raphael 2000). The lifetime prevalence rate for welfare-reliant women is estimated to be 50 to 75 percent, with a past-year prevalence of 15 to 20 percent (Bassuk et al. 1996; Colten, Cosenza, and Allard 1996; Raphael and Tolman 1997; Barusch, Taylor, and Derr 1999; Danziger et al. 2000).

Polit, London, and Martinez (2001) report 8.8 percent of women in the Urban Change sample were hit, slapped, or kicked in the year prior to the 1998 survey. Results from our longitudinal ethnographic study suggest that this may under-estimate the extent of recent or ongoing violence in this population. In our sample of 38 women in Cleveland, three women who were being abused during the course of the study did not disclose the abuse to the interviewer until after it had ended. Despite very high levels of rapport between the interviewers, these three women responded no to all questions from a self-administered questionnaire (the same one that was used in the Urban Change survey) and to open-ended questions in the interview that took place the year prior to the interview where the woman disclosed the abuse (i.e., they reported no abuse in the second year interview and disclosed in the third year interview). In each case, when they disclosed that they were being abused, the women indicated it had been going on at the

time of the prior interview and sometimes throughout the entire course of the study. We illustrate this with the case of Maria, which we have written about elsewhere (see Scott, London, and Myers 2002).

Maria had been physically abused by her boyfriend, the father of her two youngest children, for the past eight years, up until February 2000 when he moved out of the house. Maria said she was beaten at least once a month during that time, sometimes resulting in trips to the hospital for stitches. Our interviewer perceives herself to have good rapport with Maria, and this is demonstrated in the many intimate details Maria has shared about her life. Our interviewer reported being shocked that this was disclosed now. Maria has had other opportunities to talk about the abuse, but most of them while she was still living with the batterer. In the last interview, Maria was asked questions on domestic violence in our open-ended interview, as well as in a self-administered questionnaire which asks about incidents of domestic violence in the last twelve months. In both instances, she responded “no” to each question about past or present experiences of abuse.

There may be a number of reasons why Maria chose not to disclose the abuse until it had stopped, despite the high level of rapport between her and our interviewer. She may have been afraid her boyfriend would retaliate if he found out she had disclosed the abuse. She may have felt ashamed of the fact that she was being beaten, and somehow that shame was lessened when it had ended. Finally, she may have been surviving the abuse by denying either the fact of the violence or the extent of the violence. Denial becomes much more difficult if one acknowledges the violence to someone else.

Our experience with having three cases of domestic violence disclosed two years into a study involving extensive contact between interviewers and respondents surprised us. We think these data raise important methodological and policy concerns. Given that most studies cited in the literature on welfare reform and domestic violence involve single contacts, or at least significantly less contact with subjects than we have had in the Urban Change study, and given that most welfare caseworkers asking questions to identify cases of domestic violence will likely meet with recipients for much shorter periods, we have reason to be very concerned about whether we are capturing the extent of the problem either in our research or in our attempts to provide public assistance. This is particularly true for those women who are currently being abused, and those are the women about whom we should be most worried, and for whom we should seek the most flexible policies. If the shame, denial, and fear battered women typically experience keep them from disclosing an ongoing history of domestic violence to someone they have been talking with for two years, and have demonstrated otherwise a great deal of trust, why do we think this is going to be easy to identify in a welfare office in a half hour meeting? This poses an extraordinary conundrum for policy-makers.

Triangulating the Evidence: Experiments and Ethnography

In this second section of the paper, we turn our attention to two examples of combining qualitative and experimental analyses. As noted in the introduction, we agree with Gibson and Duncan (2000) that there are potentially large benefits to be gained from efforts to combine qualitative and experimental evidence. Although we did not have in place the two design features that they identify as contributing to the synergy they experienced in the New Hope Demonstration (i.e., an embedded and linked, randomly

drawn sample for the ethnography and having the same analysts conduct both qualitative and quantitative analyses), we believe that the combination of methodologies was beneficial. We begin with a discussion of the combination of qualitative and experimental findings in relation to young children's well-being in the context of maternal welfare-to-work transitions, and then we turn our attention to the same issue with respect to the well-being of adolescents.

Differences and Complementarities in Findings Derived from Experimental and Ethnographic Analyses Focused on Young Child Well-Being

Morris, Scott, and London (*forthcoming*) attempted to further our understanding of how welfare and work policies affect children by capitalizing on both the differences and the complementarities they observed in the results of two separate Next Generation reports. They sought to integrate the results into a single report in order to nuance understandings derived from either methodological approach alone. One of the initial studies synthesized findings from eight experiments (Morris et al. 2001) and the other examined longitudinal ethnographic data from Cleveland and Philadelphia (Scott, Edin, London, and Kissane 2001).

Findings from the experimental studies revealed that impacts on children's well-being differed depending on the policy approach, and subsequent changes in parents' economic outcomes. Programs that increased parents' employment and earnings benefited young children. While small, the effects were consistent enough across differing program models and samples to lend confidence regarding their replicability. By contrast, programs that simply moved mothers into employment did little to benefit young children. Because experimental impacts are derived from the *difference* in functioning between children and

parents in the program and control groups, these findings may reflect the fact that receiving welfare (as more parents in the control group are doing) and working (as more parents in the program group are doing) are equally stressful for parents. The advantages that employment may bring to parents may be outweighed by the stresses of balancing work and family. These findings point to the importance of these earnings supplement policies in increasing parents' income (as they increase employment) and in turn, benefiting children as their parents make the transition to work.

The majority of respondents in the ethnographic sample, drawn from some of the most disadvantaged neighborhoods in Cleveland and Philadelphia, did not experience stable work and sustained increases in income in this first year of the study. They struggled to make ends meet with income levels hovering around the federal poverty level. Juggling stressful jobs, irregular or part-time hours, and often temporary employment, mothers worried a great deal about their children's well-being. Exhaustion, disrupted routines, and increased absence from the home due to employment meant that some mothers reported their children were doing worse, exhibiting diminished school performance and increased problem behavior. Additionally, although perhaps less critical than these problems, they missed their mothers. However, the qualitative data also suggested that when women moved from welfare to work *and* their incomes increased, they felt their families were benefiting. This is consistent with the experimental evidence. In the qualitative sample, there were only a few women who experienced substantial increases in their incomes as a result of their labor force participation. Moreover, in addition to increased employment and income, two other factors also appeared to be critical to these families' success: 1) they had stable

employment; and 2) they felt their child care arrangements were reliable and trustworthy². Because these two factors tended to co-occur with work-related increases in income among the women in our sample, it is difficult to unpack the relative contributions of maternal employment, the income increase that accompanied the transition to employment, or employment and child care stability to the improvements in their families' lives.

As a consequence of their examination of the intersections of the experimental and qualitative results, Morris, Scott, and London (*forthcoming*) argued that these two lines of research helped them develop a more-nuanced understanding about how families fared as single parents make the transition to employment and the conditions that aided families in coping with this transition. They suggested that the qualitative findings present a slightly different account of the way in which changes in maternal employment, without increases in income, may be affecting families and children than is observed in the quantitative data.

In attempting to account for the differences they observed in the quantitative and qualitative findings, Morris, Scott, and London (*forthcoming*) noted that some conditions of employment, such as long and irregular hours, may negatively affect parents and their children, and that mandatory, rather than voluntary employment, may play out more negatively for families and children. Perhaps the fact that respondents in the qualitative study were faced with welfare time limits pressured parents to move from welfare to employment with some urgency, and increased the need to balance multiple jobs. This stress is clearly demonstrated in the ethnographic work (see also London, Scott, Edin, and Hunter 2001). In the quantitative study, parents were faced with mandates to participate

² Most of the families in our sample defined “trustworthy care” (rightly or wrongly) as kin care.

in employment. Faced with partial family sanctions may not result in the same pressure to find and keep employment as when faced with time limits on welfare benefits.

More salient, they suggested, is the fact that the qualitative findings demonstrate important countervailing influences on parents of mothers' employment not detected in the quantitative work. Predictions about the benefits of maternal employment, by enhancing maternal self esteem and increasing the regularity of routines, received virtually no support from the quantitative data. However, the qualitative work suggests that these benefits may have indeed been part of the experience of single mothers making the transition to employment. Single mothers in the qualitative work did mention increased self esteem for themselves and increased independence for their children. At the same time, however, these parents describe increased stress, disrupted routines, and exhaustion that may counterbalance these positive aspects of employment (see also London, Scott, Edin, and Hunter 2001). While the net result may be the same, understanding the effects on families as counterbalancing effects rather than neutral effects has very different policy implications.

As noted previously, unlike the quantitative findings, the ethnographies point to more costs than benefits of increases in maternal employment for children's well-being, in particular, even while the effects on parents are more mixed. What is unclear is whether such negative effects are due to differences in children's behavior that emerge over time, or real differences between groups that are attributable to differences in their employment status. In the quantitative studies, both program and control group families make the transition to employment—it is simply that a slightly greater proportion of program group members make this transition. One possibility is that effects among the

small proportion of families who were induced by the program to make the transition to employment were too small to find differences when all families are examined together. Alternatively, if the transition to employment is as stressful as remaining on welfare, no differences between groups may emerge in the quantitative data even though children may respond negatively to the transition to employment. In effect, both groups of children may be having difficulties—it may be equally difficult for children whether parents are poor and receiving welfare (and perhaps looking for work) or are working and struggling with the demands of low-wage work. Finally, parents struggling with work and family may be especially sensitive to any changes in their children’s behavior, and attribute any problems they observe to employment, rather than to other changes in their lives or developmental changes on the part of the children.

Other insights and more-nuanced understandings emerged from the qualitative-quantitative integration. The qualitative findings provided a reminder about the high levels of difficulties generally facing children and families in these studies. While not a focus of the impact estimates presented in the quantitative studies, and in some ways masked by the focus on impacts, such high levels of problems are also observed in the quantitative data. Parents in these studies report high levels of depression and stress, and children are at risk of long term difficulties. For example, about 1/3 of parents report levels of depressive symptoms that put them at risk of clinical depression, and children are performing on average at the 25th percentile on standardized tests. Even in programs that increase employment and income, and benefit children, levels of problems are still high—even after accounting for the improvements in children’s well-being observed in

these programs. This is another way that the quantitative and qualitative results are complementary.

As noted previously, the qualitative findings regarding the work-family tradeoffs women faced when they moved from welfare to work, and how such transitions affected children, are broadly consistent with the experimental findings, which show that only programs that increased maternal employment *and* income through earnings supplements improved outcomes for elementary school aged children. From the qualitative data, however, it is clear that when considering employment, income, and family well-being, we must consider a host of interacting factors. Income increases must be substantial and not offset by work expenditures; employment and income must be stable; and children must be in reliable care that is trusted by their mothers. Proponents of changes in welfare policy had suggested that one way maternal employment may benefit children is that it would increase the regularity of family routines. These findings suggest that employment itself does not necessarily lead to such regularity. But employment increases that are accompanied by increases in income, perhaps because they are accompanied by increases in stable employment and child care, are more likely to result in such benefits for family life, and, in turn, children's development.

The qualitative findings also point to the role of employment stability as an important component of the benefits of increased employment and income for young children. Interestingly, employment stability may indeed be responsible for some of the benefits of the experimental programs that increase both employment and income. Experimental findings suggest that earnings supplement programs do indeed increase employment stability as they increase the proportion of families working. While some of

these programs increased the number of jobs or spells of employment experienced by program group members (indicating a greater number of employment transitions), they all increased the length of employment spells. More specifically, the four programs that increased employment and income each increased the length of employment spells, either as measured by the length of the longest job spell (as in the New Hope program³) or as measured by the proportion of families whose first spell lasted 13 or more months (as in the other three programs). It appears that at least some of the increase in employment generated by these programs was in stable employment. As indicated by the qualitative research, this may have added some stability to these families' lives and led to some of the benefits for children.

The qualitative findings also suggest that stable and reliable child care is a key intervening mechanism in the benefits of increased employment and income. Interestingly, the quantitative studies also find evidence that programs that increase employment and income may increase formal child care and participation in after school programs (Huston and Gennetian *forthcoming*). Thus, child care, at least formal child care, may play a role in the benefits to children of earnings supplement programs. More specifically, in three of the four earnings supplement programs, mothers in the program groups were more likely to enroll their children in formal child-care programs or after-school programs and extracurricular activities than were mothers in the control group. Thus, evidence from three of the programs suggests that structured programs outside of the home may be one of the pathways by which the beneficial effects to children

³ For measures of employment stability, impacts were only presented separately for subgroups of families based on their employment status at random assignment for the New Hope program. Impacts on employment stability are only found for those sample members not employed full time at random assignment. Notably, this is also the subgroup for whom the program increased employment and income.

occurred. These findings stand in contrast to those regarding impacts of these earnings supplement programs on measures of parenting behavior. These outcomes were not much affected by these programs. Both the quantitative and qualitative findings suggest the important role child care may be playing in helping families cope with the transition to employment, and resulting in benefits to children.

Together, these findings point to the benefits to children when welfare-reliant parents increase both their employment *and* their income. Based on the quantitative findings from experimental evaluations, consistent benefits to children's achievement were observed in programs that resulted in increases in income and employment. Ethnographic findings suggest that increases in employment and income may benefit families in part because they are accompanied by stability of employment and child care arrangements. By contrast, few, and perhaps countervailing, effects were observed when parents increased their employment but not their income, or when income increases were not enough to move them substantially above the poverty threshold.

Complementarity and Hypothesis Generation in the Integration of Qualitative Data with Experimental Evidence Focused on Adolescent Well-Being

A second example of the benefits derived from having complementary qualitative and experimental evidence emerged in the Next Generation team's work on adolescent well-being (Gennetian et al. 2002). Members of the research team worked independently on analyses of the quantitative and qualitative data, but regularly held conference calls to discuss emerging findings. Case studies, developed from the longitudinal ethnographic data in Cleveland and Philadelphia, were sent as they were completed to the other team members who were working on the experimental analyses.

Consistent with the emerging experimental findings, the case studies revealed many instances of adolescents who were faring poorly as their mothers went to work. The case studies suggested that one mechanism by which adolescent school performance and other aspects of adolescent well-being might be compromised was that they were called upon to provide care to younger siblings. For example, Tina, a 35-year-old African American mother of six living in Philadelphia described the negative impact of her welfare-to-work transition on her oldest child, Tamara, as follows. Tina said that because of her work schedule, Tamara had to be responsible for waiting in the morning with the younger children until the van that took them to their day care center arrived. The van typically ran late, which made Tamara 20 to 30 minutes late for school on a regular basis. Tina said: “She’s late every day for school, every day. And what the school says to me is...they gotta do what they, what’s their policy. She’s gotta stay after school, do her detention...or she’ll lose her credit out of her, out of that morning class ‘cause she didn’t get there on time. So, she feels sad and I feel bad because I gotta be at work at 7. She can’t be at school by 7 – she can’t. We all can’t be at the same place at the same time.” Although Tina felt tremendous guilt for imposing this child care burden on her oldest daughter, she felt she had no choice. Tamara’s timely school attendance meant less for the family’s immediate well-being than Tina’s timely attendance at work. So, it was Tina’s adolescent daughter that bore the brunt of this burden in the short-term, and, potentially, in the longer-term.

Other women told similar stories, which led the group who was engaged in analyzing the experimental data to hypothesize that the negative effects they were observing might be more pronounced in those families where there were younger

children who might need care provided by an adolescent sibling. Results indicated some support for this hypothesis. The experimental analyses indicate that programs that increased employment and had negative effects on adolescents' school outcomes were the same programs that increased adolescents' home responsibilities.

Most of the experimental studies did not collect detailed information about older children's responsibilities at home. However, three studies did collect information about child care responsibilities for the subset of families that had elementary school-aged children, which allowed for an analysis of how welfare and work policies affected the proportion of adolescents who were caring for younger siblings. Rates of sibling care in these studies ranged from 14 to 28 percent at the time of the survey, with the adolescent being the primary source of care in 7 to 19 percent of cases. The majority of adolescents who provided sibling care were female (50 to 80 percent) and ranged in age from 14 to 18 years. Adolescents provided care to younger siblings for about 10 hours per week or less. In the two programs that increased maternal employment and participation in employment-related activities and negatively affected adolescents' school outcomes, adolescents were more likely to be caring for elementary-school aged siblings, suggesting that they did take on some additional responsibility at home. Adolescents in the third program, which did not adversely affect adolescent school outcomes, were no more likely than their control group counterparts to care for younger siblings. Taken together, the complementary evidence from the experiments and the ethnographic studies provides a compelling case for the role of increased home responsibilities and child care as an important mediator in the relationship between welfare and work policies aimed at adults and adolescent school outcomes.

Discussion: Models of Integration, Best Practices, and Directions for the Future

In this final section of the paper, we provide a brief and preliminary discussion of the promises and pitfalls for mixed-methods research in the context of policy evaluation research. Our comments are provisional; many are still be vetted among the co-authors and the larger group of collaborators working on the Next Generation Project.

As noted at the outset of this paper, there appears to be increasing interest in mixed-methods, qualitative-quantitative research. There are many approaches to doing this work, and substantial accumulated wisdom. Numerous recent analyses address issues related to mixed-methods research (e.g., Jessor, Colby, and Shweder 1996), including the recently-published 700-page *Handbook of Mixed Methods in Social & Behavioral Research* (Tashakkori and Teddle 2003). While a full discussion of approaches to mixed methods research is beyond the scope of the current analysis, we believe our accumulating experience suggests several important lessons about best practices and directions for the future.

Consistent with the examples and arguments put forward by Gibson and Duncan (2000), we believe it is important for there to be tighter integration of quantitative and qualitative components of studies than we have been able to achieve in the Urban Change and Next Generation projects. While our analyses have yielded valuable insights, as discussed above, the fact that there is no formal linkage between the quantitative and qualitative samples in any of the studies constrains full integration. We think there are a number of points where linkage can (and arguably should) be made that will enhance the ability of analysts to integrate the data.

First, there should be greater emphasis on designing and funding studies as mixed methods studies such that research questions and instruments are coordinated from the outset and fielded in tandem. As Gibson and Duncan (2000) note, funding constraints meant that the qualitative component of the New Hope evaluation was not fielded at the initiation of the experiment, which compromised to some extent the integration and analyses they were able to conduct.

Second, we believe that, whenever feasible, the samples should be embedded and formally linked through random selection to allow for better integration of results. This applies both to non-experimental survey research and to experiments. In combination with fielding mixed methods studies from the outset, such a linkage has benefits for analysis (as discussed by Gibson and Duncan 2000) and ensures that the time frame and location for both types of data are the same. The two examples described above, which combine experimental and ethnographic findings combine quantitative and qualitative data from different settings and time periods (i.e., different policy contexts), which necessarily weakens the formality of the integration that is possible. The Urban Change survey and ethnography were collected in the same locations and periods, but the samples were not linked, which again compromises the extent of the integration that is possible to some degree.

Third, we agree with Gibson and Duncan (2000) that having the same analysts working on the quantitative and qualitative analyses is beneficial and should be pursued whenever time, funds, expertise, and good-will for data sharing exist. Like the analysts on the New Hope evaluation, Polit, London, and Martinez (2000; 2001) conducted both quantitative and qualitative analyses themselves and found the cross-fertilization across

data sources generated new ideas and insights that might not have otherwise arisen.

Analysis of the longitudinal Urban Change survey and ethnographic data are ongoing; the extent to which systematic analysis and integration of these complementary data sources will yield important insights remains to be seen.

A fourth suggestion does not directly emerge from our experiences, and is perhaps the most novel and potentially controversial of our suggestions. However, we believe it follows logically from the considerations outlined above. We believe that there is utility in exploring and developing methods for conducting mixed methods interviews. Although there are various issues that need to be worked out and tested, we believe the integration of funding, design, data collection, sampling, and analytic personnel holds the most promise for maximizing the contributions of mixed methods research in the context of policy evaluation.

Overall, our experience suggests that mixed methods research and the integration of quantitative and qualitative research methodologies holds great promise for enhancing the evaluation of policies and programs. However, we would like to close with a caution. We must be careful not to raise expectations too high. Doing mixed methods research --- designing the study, raising the money for it, collecting the data, and analyzing it --- are difficult and require different kinds of initiatives than those that are currently pursued, funded, and staffed. Training personnel will require effort and data collection and processing is differentially expensive and time-consuming. There may be questions that one method can answer well that the other cannot, and if resources are prioritized toward one set of questions (and method) over the other, it is not clear that mixed methods research would be the best approach in that case. Moreover, inductive analyses of

qualitative data have the potential to yield valuable and unanticipated insights. However, the time involved for such analyses may be at odds with both the deductive framing and time frames for producing results that drive much policy and program analysis.

Recognizing the strengths and limitations of these methodologies, and making sure that they are not overly compromised in our efforts to integrate will be important to prevent the overselling of the promise or the under-selling of the pitfalls associated with undertaking mixed methods research.

References

- Anderson, S. A. (ed.) (1990). Core indicators of nutritional status for difficult-to-sample populations. *Journal of Nutrition*, 120, 1559-1600.
- Barusch, A., Taylor, M., and Derr, M. (1999). Understanding families with multiple barriers to self-sufficiency. Salt Lake City: University of Utah.
- Bassuk, E. L., Weinreb, L. F., Buckner, J. C., Browne, A., Salomon, A., and Bassuk, S. S. (1996). The characteristics and needs of sheltered homeless and low-income housed mothers. *Journal of the American Medical Association*, 276, 640-646.
- Brock, T., Coulton, C., London, A., Richburg-Hayes, L., Scott, E., and Verma, N. (2002). Welfare reform in Cleveland: Implementation, Effects, and Experiences of Poor Families and Neighborhoods. New York: Manpower Demonstration Research Corporation.
- Carlson, S. J., Andrews, M. S., and Bickel, G. W. (1999). Measuring food insecurity and hunger in the United States: Development of a national benchmark and prevalence estimates. *Journal of Nutrition*, 129, 510S-516S.
- Colten, M. E., Cosenza, C., and Allard, M. A. (1996). Domestic violence among Massachusetts AFDC recipients. Boston: University of Massachusetts, McCormack Institute.
- Danziger, S., Corcoran, M., Danziger, S., Heflin, C., Kalil, A., Levine, J., Rosen, D., Seefeldt, K., Siefert, K., and Tolman, R. (2000). Barriers to the employment of welfare recipients. Madison: University of Wisconsin, Institute for Research on Poverty.

- Frongillo, E. A. (1999). Validation of measures of food insecurity and hunger. *Journal of Nutrition*, 129, 506S-509S.
- Gennetian, L. A., Duncan, G. J., Knox, V. W., Vargas, W. G., Clark-Kauffman, E., and London, A. S. (2002). How welfare and work policies for parents affect adolescents: A synthesis of the research. New York: Manpower Demonstration Research Corporation.
- Gibson, C., and Duncan, G. J. (2000). Qualitative/Quantitative Synergies in a Random-Assignment Program Evaluation. Unpublished Manuscript.
- Huston, A., and Gennetian, L. A. (*forthcoming*). **Put child care paper title here**. In Larry Aber, Fred Wulczyn, and Jolyon Wurr (eds.), *New Balance: Coping with Work and Parenting in Low-Income Families*.
- Jessor, R., Colby, A., and Shweder, R. A. (eds.) (1996). *Ethnography and human development: Context and meaning in social inquiry*. Chicago: University of Chicago Press.
- London, A.S., Scott, E.K., Edin, K., and Hunter, V. (2001). *Juggling Low-Wage Work and Family Life: What Mothers Say About Their Children's Well-Being in the Context of Welfare Reform*. New York: Manpower Demonstration Research Corporation.
- Miilunpalo, S., Vuori, I., Oja, P., Pasanen, M., and Urponen, H. (1997). Self-rated health status as a health measure: The predictive value of self-reported health status on the use of physician services and on mortality in the working-age population. *Journal of Clinical Epidemiology*, 50, 517-528.

- Morris, P. A., Huston, A. C., Duncan, G. J., Crosby, D. A., and Bos, J. M. (2001). How welfare and work policies affect children: A synthesis of research. New York: Manpower Demonstration Research Corporation.
- Morris, Pamela A., Ellen K. Scott, and Andrew S. London. (Forthcoming). Effects on Children of Parents' Transitions from Welfare to Employment: Integrating Quantitative and Qualitative Research. In Larry Aber, Fred Wulczyn, and Jolyon Wurr (eds.), *New Balance: Coping with Work and Parenting in Low-Income Families*.
- Mossey, J. M., and Shapiro, E. (1982). Self-rated health: A predictor of mortality among the elderly. *American Journal of Public Health*, 72, 800-808.
- Polit, D. F., London, A. S., and Martinez, J. M. (2000). Food security and hunger in poor, mother-headed families in four U.S. cities. New York: Manpower Demonstration Research Corporation.
- Polit, D. F., London, A. S., and Martinez, J. M. (2001). The health of poor urban women: Findings from the project on devolution and urban change. New York: Manpower Demonstration Research Corporation.
- Quint, J., Edin, K., Buck, M. L., Fink, B., Padilla, Y. C., Simmons-Hewitt, O., and Valmont, M. E. (1999). Big cities and welfare reform: Early implementation and ethnographic findings from the Project on Devolution and Urban Change. New York: Manpower Demonstration Research Corporation.
- Raphael, J, and Tolman, R. M . (1997). Trapped by poverty/Trapped by abuse. Ann Arbor: University of Michigan, Research Development Center on Poverty, Risk, and Mental Health.

- Ross, C. E., and Mirowsky, J. (1995). Does employment affect health? *Journal of Health and Social Behavior*, 36, 230-242.
- Scott, E. K., Edin, K., London, A. S., and Kissane, R. J. (2001). *Unstable Work, Unstable Income: Implications for Family Well-being in the Era of Time-Limited Welfare*. New York: Manpower Demonstration Research Corporation.
- Scott, E. K., London, A. S., and Myers, N. A. (2002). Living with violence: Women's reliance on abusive men in their transitions from welfare to work. Pp. 302-316 (Chapter 14) in N. Gerstel, D. Clawson, and R. Zussman (Eds.), *Families at work: Expanding the bounds*. Nashville: Vanderbilt University Press.
- Tashakkori, A., and Teddlle, C. (eds.) (2003). *Handbook of mixed methods in social & behavioral research*. Thousand Oaks: Sage Publications.
- Tolman, R. M, and Raphael, J. (2000). A review of research on welfare and domestic violence. *Journal of Social Issues*, 56(4), 655-681.
- Zabel, J., Schwartz, S., and Donald, S. (2002). *An econometric analysis of the Self-Sufficiency Project*. Unpublished manuscript.