U.S. Department of Justice
Office of Justice Programs
National Institute of Justice

National Institute of Justice

Research Report

Project on Human Development in Chicago Neighborhoods

Technical Report I

Foreword

How do individual personalities, family relationships, school environment, and type of community interact to contribute to delinquency and criminal behavior? The Project on Human Development in Chicago Neighborhoods is an unprecedented longitudinal study that aims to answer this question by examining the development of delinquency, criminal behavior, and substance abuse from birth to young adulthood, with a particular focus on the effects of community and neighborhood contexts on individual behavior. The National Institute of Justice, in partnership with the John D. and Catherine T. MacArthur Foundation, is supporting this project to generate informed recommendations for crime prevention and intervention strategies.

163495

Project on Human Development in Chicago Neighborhoods

Felton Earls, M.D., Principal Investigator
Stephen L. Buka, Sc.D., Co-Principal Investigator

A Technical Report Presented to the National Institute of Justice

March 1997

U.S. Department of Justice
Office of Justice Programs

National Institute of Justice
Jeremy Travis
Director

Christy Visher, Ph.D.
Program Manager

Felton Earls, M.D., is a professor of human development and behavior at the Harvard School of Public Health, a professor of psychiatry at the Harvard Medical School, and Principal Investigator for the Project on Human Development in Chicago Neighborhoods.

Stephen L. Buka, Sc.D., is an assistant professor in the Department of Maternal and Child Health at the Harvard School of Public Health and Co-Principal Investigator for the Project on Human Development in Chicago Neighborhoods.

This project was supported under award number 93–IJ–CX–K005, awarded to the Project on Human Development in Chicago Neighborhoods by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. This project is being sponsored jointly with the John D. and Catherine T. MacArthur Foundation. Opinions or points of view expressed in this document are those of the authors and do not necessarily reflect the official position of the U.S. Department of Justice.

NCJ 163495

Acknowledgments

This Report reflects the efforts of many staff members of the Project on Human Development in Chicago Neighborhoods to design the components of this ambitious study and to make them operational. Special thanks is given to the following staff members who have carefully documented their work in order for these accomplishments to be shared with the research community:

Felton Earls, M.D., Principal Investigator
Stephen L. Buka, Sc.D., Co-Principal Investigator
Albert J. Reiss, Jr., Ph.D., Senior Scientific Adviser
Stephen Raudenbush, Ed.D., Scientific Director
Robert J. Sampson, Ph.D., Scientific Director
John K. Holton, Ph.D., Site Director
James R. Coldren, Jr., Ph.D., Deputy Director
Jan D. Dunham, Ph.D., Deputy Director
Patricia Lau, M.P.A., Deputy Director
Lorrie Rickman Stone, Ph.D., Deputy Director
Daniel Kindlon, Ph.D., Associate Scientist
Mary Beth Selner-O'Hagan, Ph.D., Associate Scientist
Jacqueline Barnes, Ph.D., Former Assistant Director
John D. Markovic, M.A., Research Specialist
Chris Payne, Ed.M., Senior Research Assistant
Susannah Bates, M.Ed., Special Projects Coordinator

This report was coordinated and edited by Felton Earls, M.D., and Susannah Bates, M.Ed.

Financial support for the Project on Human Development in Chicago Neighborhoods is provided by the National Institute of Justice and the John D. and Catherine T. MacArthur Foundation.

Table of Contents

Introduction

by Felton Earls, M.D., and John K. Holton, Ph.D.

The Project on Human Development in Chicago Neighborhoods (the Project) is a major interdisciplinary study aimed at deepening society's understanding of the *causes* and *pathways* of juvenile delinquency, adult crime, substance abuse, and violence. It is directed from the Harvard School of Public Health and funded by the John D. and Catherine T. MacArthur Foundation and the National Institute of Justice. (The study is described in greater detail in the booklet *Intricate Pathways*.)

The Project is unique in both size and scope, combining two studies into a single, comprehensive design. The first is an intensive study of Chicago's neighborhoods—their social, economic, organizational, political, and cultural structures, and the dynamic changes that take place within these structures. The second study will follow 7,000 randomly selected children, adolescents, and young adults, looking at the changing circumstances of their lives as well as the personal characteristics that may lead them toward or away from a variety of antisocial behaviors. The Project's interest centers on violent crime and substance abuse but also encompasses the many forms of juvenile delinquency and adult crime, from shoplifting to securities fraud. As the Project has evolved over several years of planning, its initial focus on studying problematic behaviors has broadened to include an interest in social competence. In this way the long-term objectives are relevant to creation of knowledge that not only informs violence prevention strategies but also is aimed at devising better approaches to the promotion of social competence in children from infancy to young adulthood.

Since its earliest planning meetings in 1988 to the beginning of data collection in 1994 and beyond, the Project has worked to extend the frontier of social science research in order to accomplish what may be considered the most ambitious social science research study ever attempted. Building on previous research, Project researchers have made new advances in the areas of research design and measurement, data collection, management, and analysis. The Project has taken advantage of some of the most advanced technologies available in order to facilitate and manage fieldwork activities and to manage the enormous volume of data coming in to its field office. The Project has established important contacts and agreements with agencies in Chicago in order to incorporate existing sources of data into its analytical framework. Instruments have been translated into Spanish and Polish, and bilingual research assistants have been hired to increase the Project's sensitivity in assessing diverse populations. The Project is continuously working to establish community contacts to provide accurate information about Project goals, to be responsive to public relations controversies, and to gain the long-term support and cooperation of study participants. Two newsletters have been initiated to distribute information about the study to both the local and research communities. Internal communications have also been advanced through the use of e-mail to facilitate the collaboration of staff members in different locations. A system for compiling and managing the growing collection of Project documents has been established and is being maintained. In the past year, two new components have been added to the study to examine

the earliest possible determinants of antisocial and criminal behavior by studying 6-month-old infants as well as post-traumatic stress disorder and the effects of children's exposure to violence.

Organizational Components

There are three organizational elements of the Project, which are represented in figure 0.1. First, a group of scientific directors are currently responsible for the design and execution of all components of the study. In its current form, the Project consists of an intensive study of social organization in a wide range of Chicago neighborhoods and a longitudinal study of neighborhoods in the city; a longitudinal study of several overlapping age groups to examine the types of family and individual characteristics that predispose individuals to negative outcomes of academic underachievement, criminal and violent behavior, and substance abuse; and the analytical challenge of combining these multilevel (neighborhood, family, individual), longitudinal data into an efficient program of statistical analyses. The second and third organizational elements are administrative and technical. They are meant to represent Harvard University's responsibility in supervising and managing the overall enterprise. Because Chicago is in a separate location from the university, it was necessary to replicate and extend some aspects of the university's management to this site. But the Chicago site also carries the lion's share of the scientific mission, since it is here that all the collection and initial processing of data take place.

Startup

Four interesting challenges loomed during the startup phase of the Project: recruiting and hiring staff, establishing an office culture, organizing the operational relationships, and sustaining the research mission away from the parent institution. Constructing a successful organization to accomplish the long-term tasks of the Project would be based on how well these challenges were managed.

In building a multiyear, multitask office, the first hurdle was to identify and hire staff capable of producing defensible data. The desirable candidates were those possessing a combination of complementary capabilities: exactness and flexibility, stability and adaptability, persistence and diffidence, and the ability to work independently when matched against rigorous accountability standards.

Hiring staff meant pursuing individuals who came closest to this ideal. For example, although a research assistant or field interviewer would spend up to 6 weeks learning the assessment protocol, the prototype candidate was expected to be bilingual; that is, the candidate would be proficient in a formal language such as Spanish, Polish, Russian, or Twi, and in the informal language of the *street* or a lingo specific to a subculture such as drug or cyberspace usage. The candidate also would possess a *street logic*, or sixth sense, necessary to negotiate entry into neighborhoods and to gain household access.

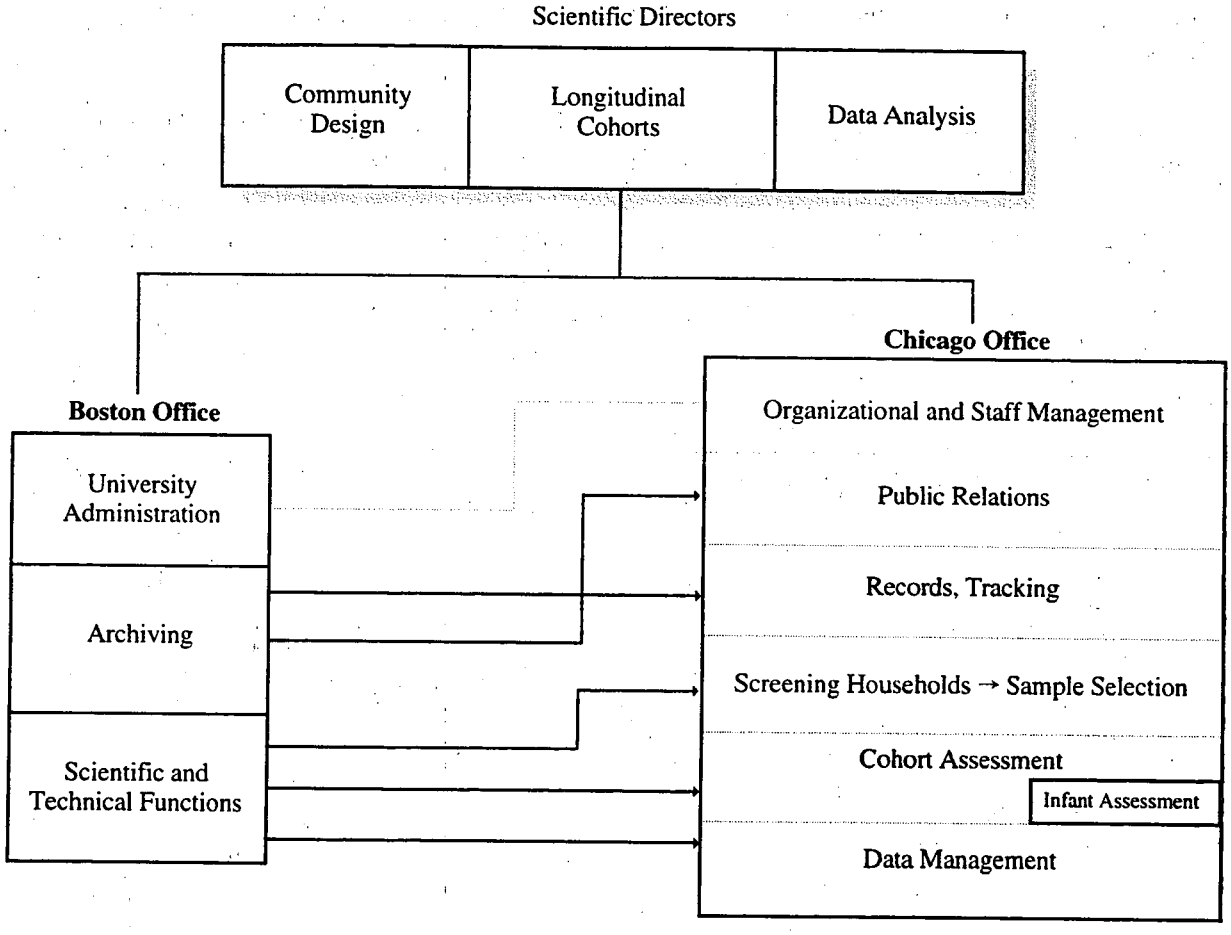
The successful interviewer would be capable of engaging potential respondents ranging from a 6-year-old to a grandparent. In a city as large and diverse as Chicago, interviewers would need to be adept at perceiving when issues of gender, race, class, and culture become paramount in their encounters. To ignore or misunderstand the subtleties of these social encounters could jeopardize the safety of an interviewer. Chicago, after all, shares a place in contemporary folklore as the home of gangsters and, although not a statistical "murder capital," still harbors the realistic threats of bodily injury and fatality within *all* of its neighborhoods.

Since these qualities were not seen as restricted to either college graduates or community activists, recruitment efforts permitted a broad range of possibilities. Although the final staff tally produced a mixture of personal employment histories, most interviewers were graduates of 4-year institutions, and many also had advanced training in neighborhood survival skills.

This technical report is a detailed account of the development of the Project's scientific products.

Figure 0.1

Project on Human Development in Chicago Neighborhoods



Part One: Survey Methods and Field Operations

Chapter 1. Sampling Plan

by Stephen Raudenbush, Ed.D., and Stephen L. Buka, Sc.D.

Study Design

The Project on Human Development in Chicago Neighborhoods combines two studies into a single, integrated design. The first is an intensive study of Chicago's neighborhoods—their social, economic, organizational, political, and cultural structures, and the dynamic changes that take place within these structures. This is accomplished through a series of data collection efforts at the community level, including a community survey of Chicago residents, interviews with neighborhood experts, systematic social observations involving block-by-block videotaping, and analyses of school, police, court, and other agency records. The second is a longitudinal cohort study involving seven randomly selected cohorts of children, adolescents, and young adults, each including approximately 1,000 participants who will be interviewed annually. The sampling plan for two key features of this design, the community survey (CS) and the longitudinal cohort study (LCS), is described in this section.

Overview

The Project's community survey is a multidimensional assessment by Chicago residents of their neighborhoods. Neighborhoods are operationally defined as 343 clusters of city blocks, hereafter called "neighborhood clusters" (NCs), chosen to be geographically compact and internally homogenous with respect to socioeconomic and ethnic mix, housing density, and family structure.

The 343 NCs contain all of the dwelling units in Chicago, and every NC was represented in the sample. Thus, the NCs are strata rather than sampling units for purposes of the CS. The basic design for the CS has three stages: at stage 1, city blocks were sampled within each NC; at stage 2, dwelling units were sampled within blocks; and at stage 3, one adult resident was sampled within each selected dwelling unit.

Although this three-stage design within each stratum was straightforward, the sample size and method of sampling differed depending upon whether or not an NC was one of the 80 NCs selected for the Project's LCS. The 80 selected NCs are referred to as "sample NCs," and the remaining 263 NCs are designated as "nonsample NCs" (all 343 are represented in the CS).

The target sample size for each nonsample NC was 20. The CS sampling plan for the nonsample NCs called for selection of nine blocks within an NC, three dwelling units within a block, and one resident within a dwelling unit, potentially producing 27 (9 x 3) interviews. Given a dwelling unit occupancy rate of .90 and a response rate of .85, about 20 completed interviews per nonsample NC could be expected.

For the 80 sample NCs, the target sample size was 50. The CS sampling plan for the sample NCs built upon work already done in selecting blocks for the LCS. Within these blocks, a systematic random sample of, on average, 65.4 dwelling units was selected, with one respondent sampled per dwelling unit. Again, given a .90 dwelling unit occupancy rate and a response rate of .85, this sample was expected to yield about 50 completed interviews per sample NC.

In sum, the sample design for the CS called for 20 interviews per nonsample NC, yielding 5,260 (20 x 263) completed interviews from the nonsample NCs, and 50 interviews per sample NC, yielding 4,000 (50 x 80) completed interviews from the sample NCs, for an overall target sample size of 5,260 + 4,000 = 9,260. To obtain a more detailed understanding of the sampling plan requires discussion of (a) the construction of the 343 NCs, and (b) sample selection within the sample and nonsample NCs.

Construction of 343 NCs

Chicago's 847 populated census tracts were combined to create 343 NCs. The overriding considerations in forming NCs were that they should be ecologically meaningful areal units composed of geographically contiguous census tracts, internally homogeneous on a variety of census indicators, and sufficiently large to yield adequate and nearly balanced sample sizes of persons for the LCS. A series of cluster analyses of census data helped guide the construction of internally homogenous NCs with respect to racial-ethnic mix, socioeconomic status, housing density, and family organization. Random effects analyses of variance produced intracluster correlation coefficients to assess the degree to which this goal had been achieved; the analyses revealed that the clustering had been quite successful in producing internal homogeneity of NCs. Geographic boundaries (e.g., railroad tracks and freeways) and knowledge of traditional local community areas also were used to construct NCs with ecological integrity.

Census data were used to operationalize the two stratification variables: racial-ethnic mix (seven levels) and socioeconomic status (three levels). The NCs were cross-classified by these 2 variables, as displayed in table 1.1, and a stratified probability sample of 80 NCs was drawn for the LCS.

Table 1.1: Final Racial/Ethnic Composition by Socioeconomic Status (SES) Strata (NC = 343)

Racial/Ethnic Strata	Low SES	Medium SES	High SES	All SES Groups
75% or More African American	77	37	11	125
75% or More White	0	5	69	74
75% or More Hispanic	12	9	0	21
20% or More Hispanic/20% or More White	6	40	12	58
20% or More Hispanic/20% or More African American	9	4	0	13
20% or More African American/ 20% or More White	2	4	11	17
Neighborhood Clusters Not Classified Above	8	15	12	35
All NCs	114	114	115	343

Selection of NCs for the LCS. The number of NCs falling into the 21 strata created by the cross-classification of racial-ethnic mix and socioeconomic status was quite variable. The aim of the Project was to obtain nearly equal numbers of NCs from each of the strata. In fact, 3 of the 21 strata were empty, and an additional 3 cells had fewer than 5 NCs. NCs were selected with certainty in these 3 cells. In other strata, 4 NCs were selected by means of a

systematic random sample after sorting by socioeconomic status and housing density.[1] Because of the sparseness of several strata, a sample of 4 NCs per stratum could not produce the desired total sample of 80 NCs. The balance of NCs were sampled from the largest strata by means of a systematic random sample.

CS Respondent Selection Within Nonsample NCs

As mentioned earlier, the aim was to select nine blocks per NC and, within each block, three dwelling units, with one respondent per dwelling unit. Blocks were selected with probability proportional to size; dwelling units within blocks and persons within dwelling units were selected at random. The tasks at each stage are described below.

Selecting blocks. The method of selection was a systematic random sample with probability proportional to the number of dwelling units in the block. To achieve this, blocks were listed in block number order. Associated with each block was the number of dwelling units in that block as given in the 1990 census, along with the cumulative total number of dwelling units. Nine equal intervals between α and $N + \alpha$ were computed, where N is the number of dwelling units in the NC and α is a random number between 0 and N/9. Associated with the endpoint of each interval was a particular block, and that block was selected into the sample.[2]

Selecting dwelling units. All of the dwelling units within a selected block were listed, and, from the list, three were selected at random.

Selecting respondents. Within each selected dwelling unit, a list of persons 18 years and older was obtained, and one person from that list was selected at random to be interviewed.

CS Respondent Selection Within Sample NCs

For sample NCs, the goal was to interview 50 survey respondents. Under assumptions of an occupancy rate of .90 and a response rate of .85, 65.4 survey respondents per NC were required.

Selecting and listing of blocks. To accomplish the purposes of the LCS, a large number of blocks were selected and listed as described in detail elsewhere. These blocks constituted a simple random sample of all blocks in the NC.

---

[1]The advantages of systematic sampling for increasing precision and reducing costs are discussed, for example, by Scheaffer, Mendenhall, and Ott, *Elementary Survey Sampling*, Boston: Duxbury Press, 1990, chapter 7.

[2]If a block contained more than N/9 dwelling units, that block was selected with certainty and removed from the list. A new interval was then constructed having width N/8 and a systematic random sample of size 8 was drawn. Blocks with fewer than three dwelling units were treated as if they had three dwelling units.

Selecting dwelling units. A systematic random sample of dwelling units was drawn within the set of listed blocks in each sample NC. First, dwelling units were sorted by census tract, by block group within census tract, by block within block group, and by address within block. Equal intervals of length K/65.4 between β and K + β were computed, where K is the number of addresses in the NC and β is a random number between 0 and K/65.4. Associated with the endpoint of each interval was a particular address, and that address was selected into the sample.

Selecting respondents. As with nonsample NCs, a list of persons 18 years and older was obtained within each selected dwelling unit, and one person from that list was selected at random to be interviewed.

Selection of Subjects for the Longitudinal Cohort Study

Dwelling unit listing. For each of the 80 target NCs, dwelling units were listed and sampled for subsequent household member enumeration and subject enrollment. This entire process was completed on a neighborhood-by-neighborhood basis. First, dwelling units within an NC were listed using established field procedures. In most instances, all units within an NC were listed; for particularly large NCs, a random sample of blocks was selected for listing (proportional to the estimated number of eligible subjects within the NC). Within each NC, the listed dwelling units were assigned at random to 1 of 12 equal-sized replicates. Replicates were consecutively assigned for enumeration and subject enrollment, with all dwelling units per replicate released for fieldwork until the requisite sample size was drawn.

Household enumeration and subject enrollment. All sampled dwelling units were enumerated using a specially designed survey instrument that ascertains the ages and genders of all household members. All household members who are within 6 months of one of the seven cohort ages at the time of enumeration are eligible for study participation (e.g., 2 years 6 months to 3 years 6 months; 17 years 6 months to 18 years 6 months). All pregnancies are treated as eligible for study participation. The recruitment of all eligible participants within a household ensures that the sample will reflect the household composition of the NC for each cohort and minimizes the variance of estimated NC-level parameter estimates.

Chapter 2. Sample Retention and Locating

by James R. Coldren, Jr., Ph.D.

A sample retention and locating plan was developed and is coordinated by the Agency Records/Sample Retention Unit (AR/SRU) to guide this task over the course of the study. Often referred to as "subject tracking," the Project's approach to sample retention and locating is governed by the following principles and assumptions:

- Sample retention and locating is a projectwide responsibility. Success in sample retention and locating participants will depend on the coordinated efforts of all major units within the Project.

- Sample retention is a proactive task. Rather than waiting for study participants to become hard to find, participants should be contacted periodically and made to feel linked to the Project.

- Guidelines and principles for sample retention and locating should be established immediately. They will come into play during Wave I (for eligible participants, identified by Abt Associates, who are hard to locate), and more so during Waves II–VIII.

- Sample retention and locating procedures will emerge over the next few years as assessment Waves I–III are completed and as methods are crafted. Immediate tasks and responsibilities must be established, but these will change as project experience is accumulated.

The Project's approach to sample retention and locating is explained in more detail below and is followed by a preliminary model for participant locating, a three-stage process of increasing agency effort that will be implemented once a study participant becomes hard to find.

Sample retention and locating tasks are not well defined in the literature and the Project's work is so novel, Project staff will have to develop standards and procedures. For these reasons, the locating model will be evaluated continually. Changes to the model, procedures, and responsibilities will be made as necessary; this will require routine feedback via debriefings of field staff and supervisors. Included is a checklist of questions, statements, and observations to be made during the assessment that will assist in locating participants in the future and in assessing risk for attrition. The final section presents a plan of action (list of tasks) to establish the sample retention function.

Approach to Sample Retention and Locating

The goal of sample retention is twofold: (1) to keep as many study participants as possible engaged in the project, and (2) to collect as much information and develop as many resources as possible for use in locating study participants who move or otherwise become hard to find.

Sample retention will entail locating procedures, and at times staff interviewers will have to use street knowledge and every means possible to locate study participants who are particularly hard to reach. The work of sample retention, however, is up-front and preventive in nature—the goal is to prevent subject attrition. The sample retention task, then, becomes one of using early contacts with all study participants—the screening process conducted by Abt Associates, the enrollment process, and Wave I assessments conducted by Cohort Assessment Unit (CAU) research assistants—as the place to begin the work of sample retention. Recording data from these early contacts and periodic contacts between assessment visits will work to the Project's benefit, as will reminding participants of their involvement with the Project, of the Project's need for them in the future, and of their importance.

This approach to sample retention suggests that much of the work will be done proactively. Interviewers will capitalize on the opportunities provided by early contacts by instilling pride and commitment to the Project and by assessing attrition risk, rather than waiting for people to disappear and then tracking them down (see chart 2.1).

Early contacts with study participants provide opportunities to emphasize their importance to the Project, to make observations and collect information pertinent to the future task of finding study participants, and to ask pertinent questions about their future plans. Information gained from Abt Associates, the Wave I assessments, and agency records will be used to locate people during assessment Waves II–VIII. Plans should be developed to contact longitudinal and community survey respondents during interim time periods and to develop other incentives for individuals to participate in the Project. These plans may include an ongoing mail campaign, identifying those at risk for attrition and focusing contact efforts on them, and periodic mailing of newsletters or other notices.

A proactive approach to sample retention should include incentives for participants other than payment for their time. Inexpensive gifts for participants should be considered, as should gifts to individuals and agencies who are helpful with locating tasks. Other projects have used "finding fees," paying individuals for locating participants or for information, or donations (such as books for school libraries). These ideas and options should be given immediate consideration. They may involve devoting staff time to developing corporate and foundation relations.

A 3-Stage Model for Locating Study Participants

Some participants will be hard to locate or will simply drop out of the sample from time to time (for varying amounts of time).[1] To make the best use of project resources and apply them to the task of locating participants, a three-stage model for locating participants will be used. Chart 2.2 depicts the three-stage locating process.

Stage 1: CAU Locating Efforts Using Inhouse Resources

As the Project's primary field staff, research assistants and supervisors will make the initial attempts to locate longitudinal cohort members to schedule and conduct assessments. They constitute the front line of sample retention and locating. When they determine that a longitudinal cohort member is hard to find or has moved, they will consult inhouse data resources in an attempt to locate the person, or they may refer the locating task to the Agency Records Unit (ARU). The Project's data management system (DMS) will have special reports with pertinent locating information (e.g., names and addresses of friends and relatives, school information, work information) that staff may request. Once these remedies are exhausted, a more intensive search of records will take place.[2]

Stage 2: Locating with Agency Records and Contacts

At this point, ARU should be involved in the locating task. When the locating resources in Stage 1 do not result in a find, a more intensive search into inhouse agency record data files will be undertaken, and ARU staff may request special record searches at any of a variety of agencies, both in and out of State. Visits to these agencies (schools, jails, hospitals, etc.) may be required.

Stage 3: Intensive Locating

This is the most intensive search process. ARU may rely on other online searches, some of which involve a fee. Within reasonable geographic limits, the ARU sample locator will work with CAU staff to determine the best places to look for a participant and use every means

---

[1]If, for example, 30 percent of 24,000 respondents (15,000 participants plus primary caregivers, plus 9,000 community survey respondents) will move (or be hard to locate for other reasons) over the next 8 years, and moves will take place about every 2.5 years (some will be out of State), then approximately 3,000 participants, or approximately 250 per month, will move or be hard to locate each year.

[2]The point at which a research assistant or supervisor requests locating assistance from ARU is not clearly defined, nor should it be at this early stage. It will depend on the current caseload as well as on the status of DMS development (availability of online locating resources) and other factors. Locating responsibilities regarding community survey respondents have not been addressed. It is unclear how the responsibility for this task will be shared between Project and future survey contractors for the community household surveys and key informant interviews.

possible to locate the person and arrange an interview. Visits will typically be made to the neighborhood, schools, workplaces, churches, local stores, and businesses.

It is important to view this process as an all-out team effort. Successful retention and locating will depend on the skills and resources in each of the Project units, as explained below.

Sample Retention and Locating Responsibilities

Sample retention is a projectwide responsibility with shared tasks and responsibilities across all units, coordinated by AR/SRU.

Cohort Assessment Unit. The initial and annual contacts with study participants by CAU field staff represent critical moments for engaging participants for the long term, impressing upon them the importance of their participation, conveying caring messages, and assessing risk for attrition or moving. ARU will depend on CAU for risk assessments, warnings about potential or actual loss of participants, and periodic debriefings on field experience (Abt Associates and National Opinion Research Council interviewers should be debriefed as well). Most participants should be found during Stage 1 in the locating process. With work and caseload demands as they are, successful locating in Stage 1 will depend on fluid communication between CAU and ARU.

Data Management and Quality Control Unit (DMQC). DMQC will provide data and computing services needed to conduct risk assessment and locating. Early warning/risk indicators may come from the Abt Associates screening data and from enrollment/Wave I data. A special set of computer files or tables may be created to hold contact/locating records. Standardized reports will be created and accessible on DMS. Since locating will rely on agency and Project cohort assessment records, links with DMQC will be constant. Locating some participants will require modem communications and access to external information services, which also will require support from DMQC.

External Relations Unit. Good community relations will be an essential component of sample retention and participant locating activities. Community organizations may be contacted as the Project looks for study participants or for advice regarding how to locate individuals—tasks that require trusting relationships between the Project and the community. To the extent that Project staff develop corporate relations to secure donations to obtain other incentives for individuals and agencies, the External Relations Unit will play a key role.

Board of Scientific Directors. There will be close oversight of the sample retention/attrition problem. Expectations, tolerances, and resources relating to sample retention and locating should be clearly communicated, and Project activities in these areas will be monitored closely.

Early Assessment of Risk for Attrition

This section presents a form, the Checklist for the Early Assessment of Risk for Attrition (see checklist 2.1), that CAU research assistants will complete at the end of each assessment, along with the Project release of information form. The checklist covers nine items that field interviewers should either ask respondents, make observations about, or mention to respondents at the end of the assessment process. This information will contribute proactively to sample retention by producing information that will help identify those at risk for attrition and locate participants in the future.

Plan of Action

Table 2.1 presents an action plan for initiating sample retention and locating activities. It identifies activities at three levels of urgency—immediate (starting December 1994), short term (through January 1995), and long term. Pending review and discussion, these activities (and any modifications) will be incorporated into the ARU management and intelligence plans.

Chart 2.1
Opportunities for Sample Retention Activities

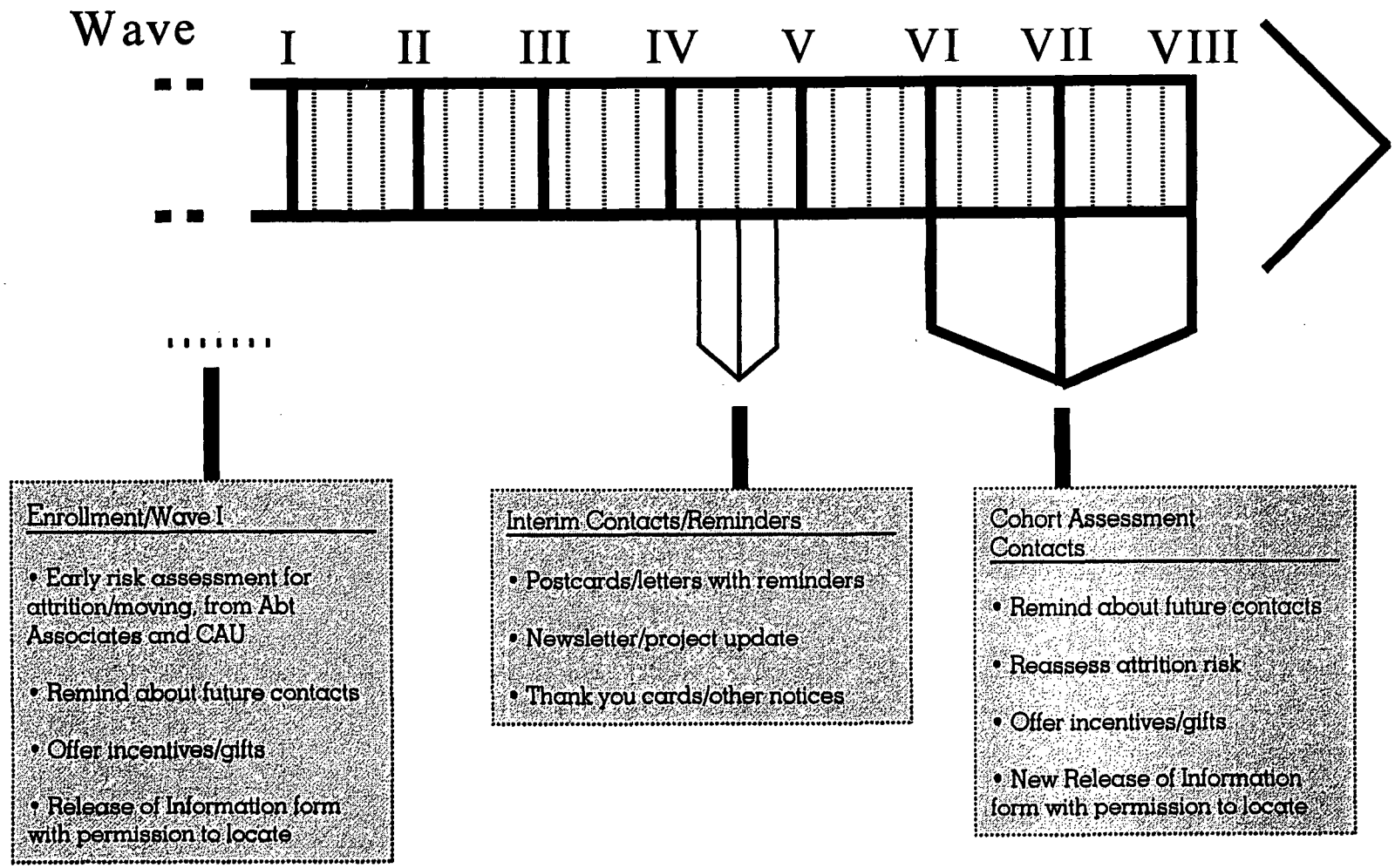
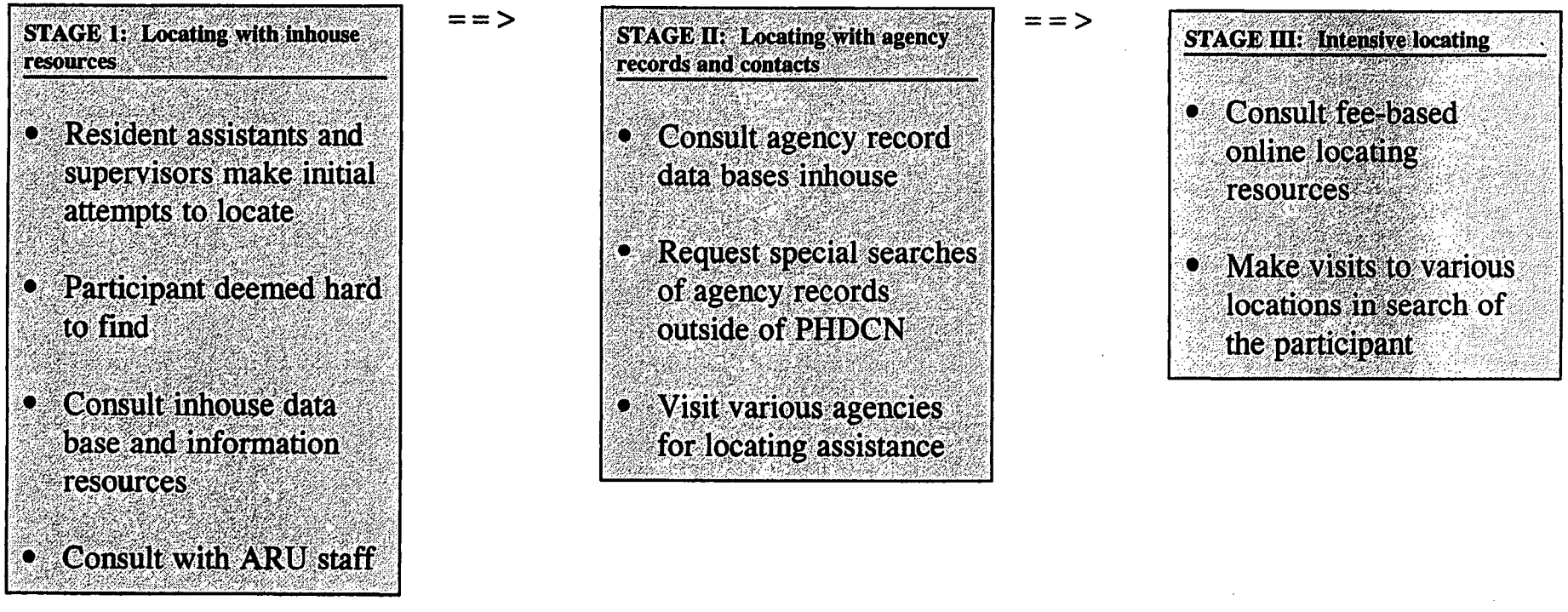


Chart 2.2
Preliminary 3-Stage Model for Sample Locating



Checklist 2.1
Project on Human Development in Chicago Neighborhoods
Checklist for the Early Assessment of Risk for Attrition

Ask questions 1–4 of the respondent.

1. Are you planning on moving/leaving home in the next year? Y N

If "Yes," where (city/State/neighborhood)?

_____

When? _____

2. Do you plan on changing jobs in the near future? Y N

If "Yes," when? _____ New employer? _____

New employment address: _____

_____

3. Do you plan on changing schools or going away to school? Y N

If "Yes," what will the new school be? _____

New school address: _____

_____

4. Can you give me the names, addresses, and phone numbers of two close friends or relatives who might be able to help us locate you in the future?

Name: 1._____ 2._____

Relationship: _____ _____

Address: _____ _____

_____ _____

City/State/Zip: _____ _____

Phone: _____ _____

Try to make the following observations at the close of the session.

5. Make note of any observations or statements made by the respondent that suggest he/she may be moving or does not consider his/her current housing permanent.

_____

_____

6. Did the respondent seem (check one):

_____ annoyed _____ nervous _____ bored _____ afraid

Please note any other observations about the respondent's demeanor: _____

_____

7. Using the scale below, make your best assessment of the respondent's risk for attrition:

Circle one:

1	2	3	4	5
Least risk for attrition			Greatest risk for attrition	

Making statements along the lines suggested below will help prepare respondents for future assessments.

8. Let the respondents know that we will be contacting them in the future to arrange the next visit and that they may be contacted by phone or mail in the interim.

9. Let the respondents know how important they are to us, how what they tell us will be helpful to many other families, how you will look forward to seeing them/their children grow over the years, etc.

Table 2.1

Sample Retention and Locating Activities		
Immediate (starting December 1994)	**Short Term (through January 1995)**	**Long Term**
Hire the sample retention coordinator	Visit with Northwestern University longitudinal study team	Implement interim contact plans for community and longitudinal cohorts
Work with CAU to develop guidelines and principles for sample retention and locating	Interview and hire sample locator	Additional programming and technical requests to DMQC
Establish a schedule for debriefing Abt Associates and Project interviewers	Conduct in-service with CAU	Continue staff in-service
Establish a schedule for accompanying Project on Human Development in Chicago Neighborhoods (PHDCN) RAs into the field	Monitor and revise sample retention plans	Develop PHDCN risk model for attrition
Revise the ARU management and intelligence plans	Prepare written procedures and principles for sample retention and locating activities	Solicit corporate donations for incentives and gifts
Continue research- and information-gathering activities relating to sample retention	Develop incentives and interim contact ideas	Develop periodic attrition/locating report formats for the Board of Scientific Directors
Reassess agency records access with sample retention in mind	Develop a corporate donation plan for gifts and incentives	Develop out-of-State connections for locating purposes
	Obtain an account number from the post office for mailing indicia	Revise the PHDCN study participant contact log and relevant portions of DMS
	Conduct research on mobility within and out of Chicago	Work with DMQC to develop standard locating screens and reports
	Establish contacts with locating/reference services	

Chapter 3. Computer Mapping Technology

by James R. Coldren, Jr., Ph.D., and John D. Markovic, M.A.

As PC-based computer mapping technology has become increasingly advanced, user friendly, and affordable, the Agency Records and Sample Retention Unit (AR/SRU), led by James (Chip) Coldren, has taken advantage of the opportunity to develop its own sophisticated mapping capability, which now plays an important role in the Project. There are three basic types of maps that offer a unique way to illustrate and examine information:

- *Boundary* or area maps show the geographic borders of "recognized areas."
- *Thematic maps* use different shades, patterns, or colors to display variation in a measure across geographic areas.
- *Pin-maps* display the location of events, buildings, organizations, etc.

These different types of maps can also be combined to illustrate even more powerful information. AR/SRU is currently using computer mapping technology in several ways to support the Project's goals.

Planning for Field Work

AR/SRU creates maps of various types as aids for supervisors and research assistants in the field as well as for Project administrators. Frequently, AR/SRU creates maps depicting small areas (most often a single Project neighborhood cluster) with cross streets, census tracts, and blocks identified. These maps are helpful to field supervisors who generally assign household screening tasks by census tract and block. One of the more popular maps to date displays the entire city with Project neighborhood clusters and major cross-streets identified; this map is referred to often whenever cluster-oriented planning is taking place. Another map uses U.S. census data to identify areas with high concentrations of linguistic isolation, assisting field supervisors in making field assignments for multilingual staff.

Identifying Boundaries Within Chicago

Neighborhood cluster boundaries are of great interest to the Project staff, but there are a host of other boundary issues with which to contend. For example, the Project clusters are nested within the 77 commonly recognized Chicago community areas (in a few instances, the Project clusters cross community area boundaries). AR/SRU is in the process of creating a series of boundary maps that clearly display the relationship of the neighborhood cluster boundaries with other political or administrative boundaries such as police districts and beats, ZIP code areas, and school boundaries. As the research plan unfolds, the relationship of Project neighborhood clusters to other political or administrative boundaries will be examined with the help of computer mapping.

Studying the Chicago Context With Maps

A key component of the research agenda in Chicago includes studying citywide context from various perspectives and using different data sources. AR/SRU is regularly seeking out sources of geo-based information (information that can be linked, by address, to coordinates or with specified boundaries on a map) to examine social, political, economic, or other important phenomena across the city. AR/SRU has created thematic maps depicting poverty or crime across census tracts in Chicago, the distribution of scattered public housing, deaths related to the recent heat wave (see map 3.1), and the like. Such maps typically attract interest because they are cognitively appealing to audiences and they have the power to communicate effectively what used to be attempted with large tables of numbers. Maps also represent the beginning of an important research tradition at the Project—studying and depicting the distribution of aggregate measures of neighborhood quality of life across space (neighborhoods) and time.

Recently, the Project created several maps illustrating: (1) the geographic convergence of crime and low birth weight (see map 3.2), and (2) the geographic convergence of crime and educational attainment (see map 3.3).

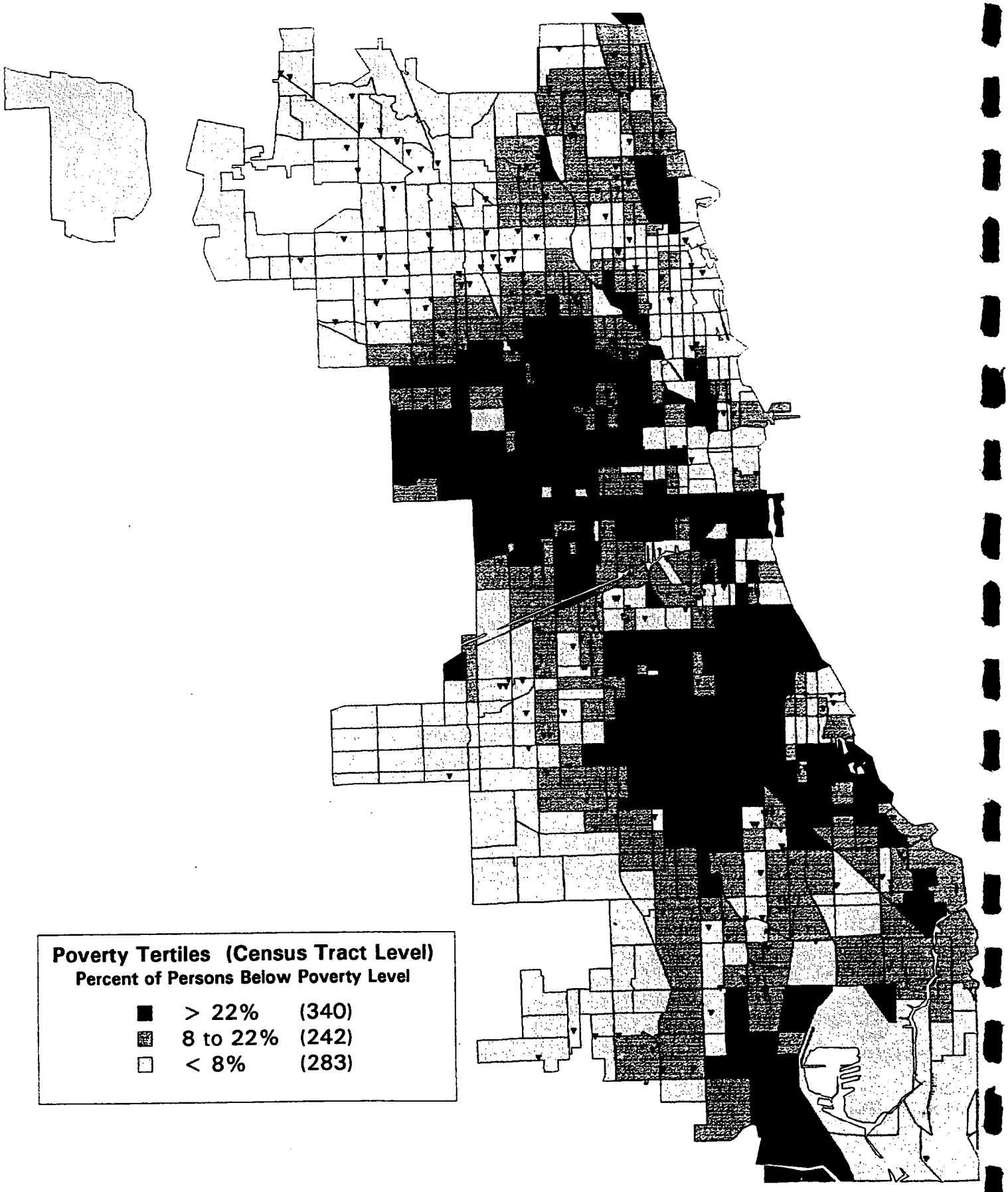
In addition, AR/SRU is laying the groundwork for the study's future use of mapping technology. Because the project will be dealing simultaneously with a variety of community definitions, data will be maintained at various levels of aggregation. For example, mobility data from the 1990 census are available at the census tract level and can be aggregated easily to the neighborhood cluster and community area levels (since census tracts are contained entirely within these larger geographic units). Some types of aggregation will present unique challenges. For instance, attributing school performance scores to neighborhood clusters will be no simple matter since school boundaries sometimes cut across neighborhood cluster boundaries. If a particular neighborhood cluster is served by three elementary schools, it may be necessary to explore techniques that account proportionately for each school's contribution to a neighborhood cluster score. Therefore, the development of the Project's computer mapping capabilities must proceed in a deliberate and careful manner.

The possibilities for new developments in computer mapping lead down some exciting and challenging paths. The Project community survey that was recently completed contains information pertaining to "cognitive boundaries," how individuals define their neighborhoods in terms of space (north, south, east, and west boundaries). Comparing self-reported neighborhood boundaries to more arbitrary and official definitions should produce interesting results, as should the comparison of self-reported boundaries across individuals by age. As the schools, police, juvenile justice, public housing, and other major social institutions in Chicago undergo massive planned changes over the next decade, computer mapping will help in describing and evaluating how those changes are reflected across different sectors of the city.

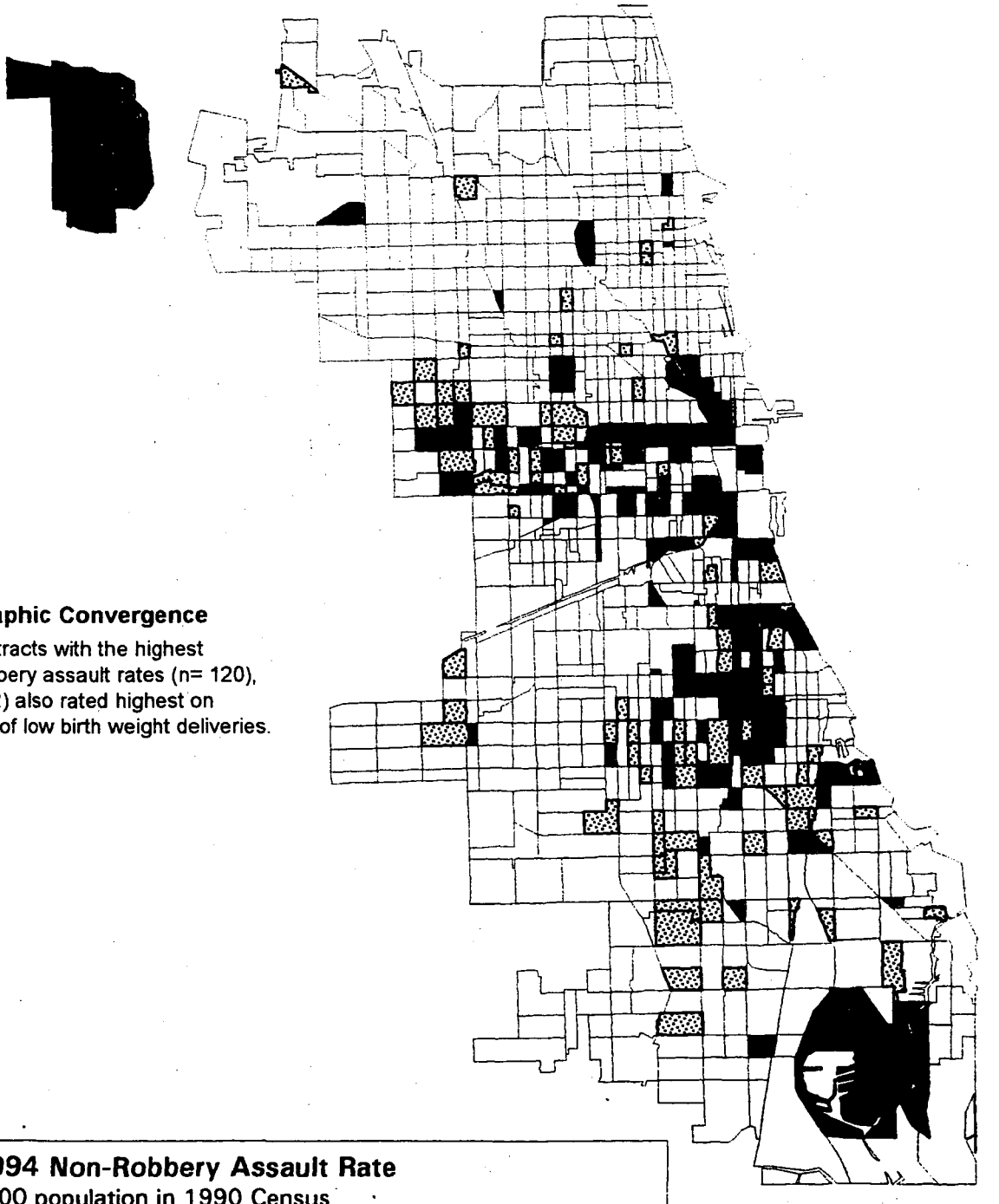
The power of computer mapping lies more in the effective communication of data than in data analysis. As the Project mines its vast data resources from both interviews and agency records, computer mapping will be an essential vehicle for presenting the Project's findings to different audiences.

MAP 3.1

Identified Victims of Chicago's 1995 Heat Wave and Census Tract Poverty Level (1990)



Geographic Convergence of Crime and Low Birth Weight



Geographic Convergence

Among tracts with the highest non-robbery assault rates (n= 120), 35% (42) also rated highest on percent of low birth weight deliveries.

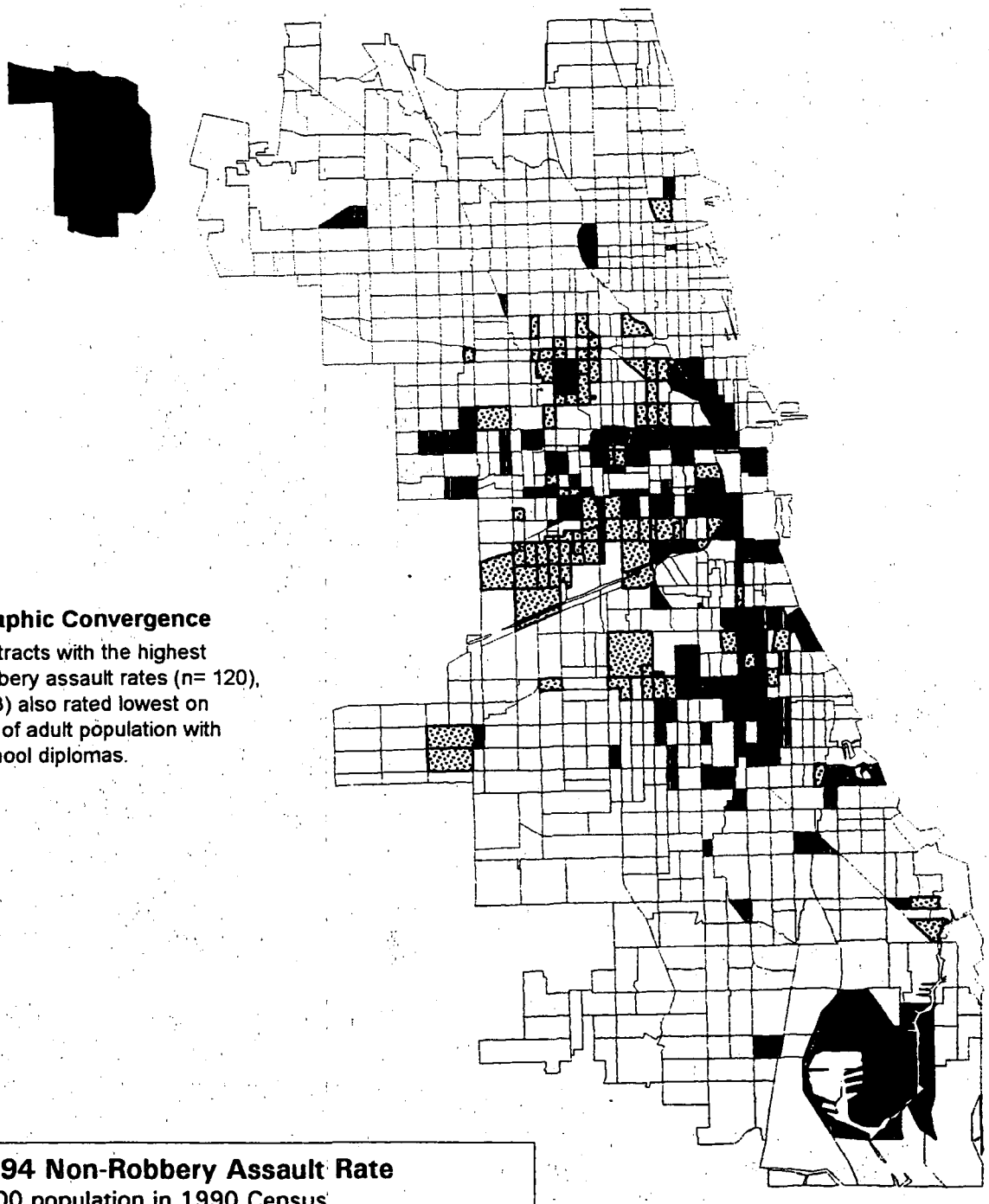
Highest 1994 Non-Robbery Assault Rate
per 1000 population in 1990 Census

■ > = 100 (120)

Highest Percentage of Low Birth Weight Deliveries
per total 1994 births (location determined by census tract of mother's residence)

▩ > = 18% (130)

Geographic Convergence of Crime and Educational Attainment



Geographic Convergence

Among tracts with the highest non-robbery assault rates (n= 120), 28% (33) also rated lowest on percent of adult population with high school diplomas.

Highest 1994 Non-Robbery Assault Rate
per 1000 population in 1990 Census

■ >= 100 (120)

Lowest Percentage of High School Graduates
per adult popultion (> = 18) in 1990 Census

▩ <= 42% (122)

Chapter 4. Accessing and Assessing Agency Records

by James R. Coldren, Jr., Ph.D., and John D. Markovic, M.A.

The primary tasks of the Agency Records and Sample Retention Unit (AR/SRU) are to:

- Identify, document, and assess the quality and availability of data obtainable from public and private agencies that are pertinent to the Project.
- Develop working agreements with those agencies.
- Collect data for Project research and management purposes from agency files and data bases.

These data generally include the outcome measures of delinquency, criminality, school performance, and various health indicators; aggregate measures of these and other community characteristics (offending rates, community structure, birth rates, and other quality-of-life indicators); and data that generally describe study participants' contacts with public (chiefly governmental) agencies in Chicago, Cook County, and the State of Illinois. AR/SRU may be the first formalized department within a large-scale research project to focus on issues of agency records, and it has carefully designed, implemented, and documented the process used for accomplishing these goals. AR/SRU prepared a paper titled "A Systematic Process of Accessing and Assessing Agency Records for Longitudinal Research" (Coldren, Markovic, McElvain, and Lazer, 1995) to document the details of the process and contribute to the field of research management by explaining operational details, or how things get done. This paper is, in part, summarized below.

The primary use for agency records will be to obtain outcome measures, particularly of youth and adult offenders, from self-report instruments and, to the extent possible, to document careers of offenders (including onset, persistence, desistence, seriousness, frequency, and sanctions) using official data from police, courts, and corrections agencies. School performance is another outcome measure that agency record data will provide.

Other important uses for agency record data include collecting current and historical data at aggregate levels (census tract, schools, school district, police district, community area, etc.) and collecting data that can be used to validate or supplement Project measures at the individual or community level. Aggregate data can be used to develop historical, social, and political contexts for the Project as well as to develop community quality-of-life indicators to correlate with the Project's various community measures.

As discussed in chapter 2, "Sample Retention and Locating," AR/SRU also anticipates relying on agency record data to locate study participants who are hard to find. It is not known at this juncture how difficult it will be to locate participants in subsequent years, but this is a task for which agency record data will be helpful.

To collect record data, AR/SRU has established working agreements with the following nine agencies: Chicago Police Department, Illinois State Police, Juvenile Court of Cook County, Cook County Criminal Court (including Adult Probation), Cook County Department of Corrections, Illinois Department of Corrections, Chicago Board of Education (public schools), Archdiocese Office of Catholic Education (Catholic schools), and the Chicago Department of Health. Table 4.1 lists the types of data being collected by each of the agencies. AR/SRU expects to develop agreements with other city agencies, suburban police agencies and courts, and child and public welfare agencies over the next several years.

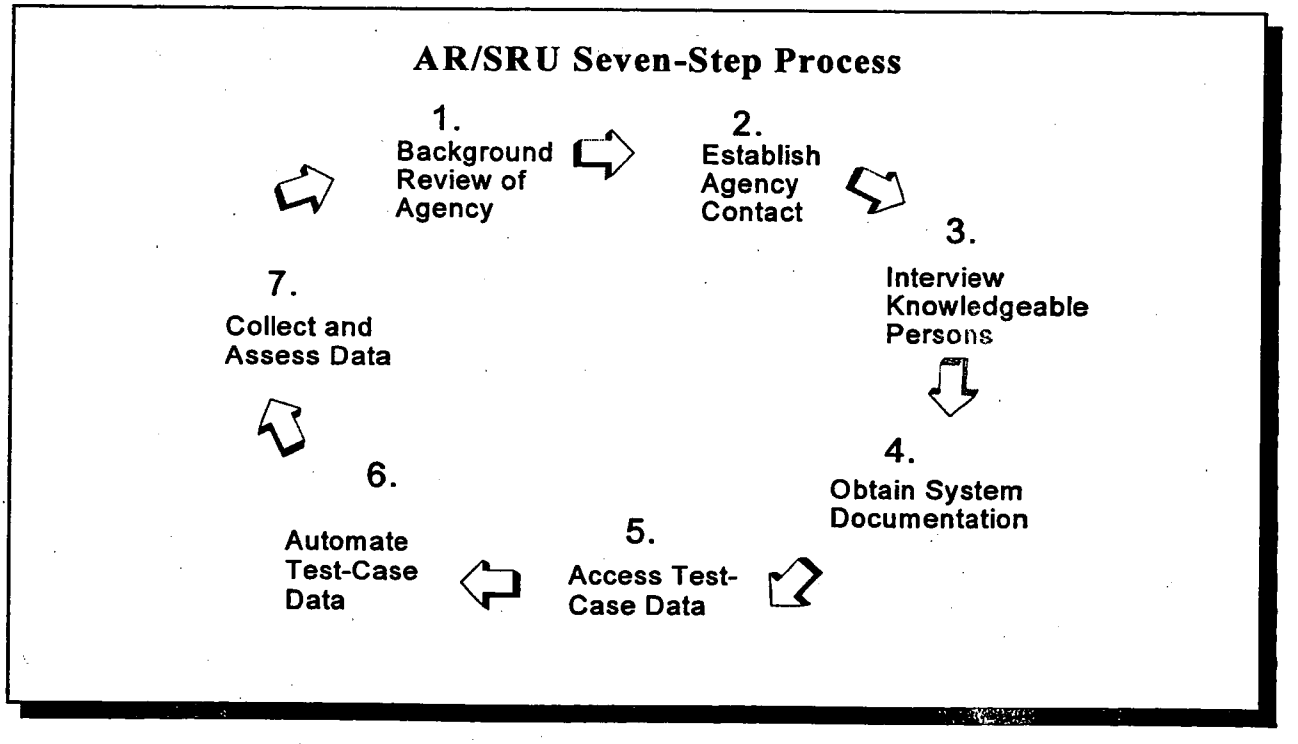
Table 4.1
AR/SRU Activity with City, County, and State Agencies

Agency	Types of Data
Chicago Police Department	Individual youth community adjustment and arrest records, adult arrest records, police dispositions, aggregated incidents (reported offenses) at the census tract level
Illinois State Police	Adult arrest and criminal history records
Juvenile Court of Cook County	Court dispositions for youth arrests, informal dispositions, sentencing outcomes, youth detention
Cook County Criminal Court	Court dispositions for adult arrests and for youth arrests transferred to adult court, dispositions, sentencing outcomes
Cook County Department of Corrections	Time spent in detention for arrests, time served for minor offenses and postrelease violations
Illinois Department of Corrections	Time served in prison, postrelease violations
Chicago Board of Education	School enrollment and transfers, attendance, standardized test scores, participation in special education, passing or being held back, exit reasons; school-level climate indicators such as attendance, percent low income, mobility, chronic truancy, etc.
Archdiocese of Chicago, Office of Catholic Education	School enrollment and transfers, attendance, standardized test scores, passing or being held back, exit reasons; school-level climate
Chicago Department of Health	Birth and death records aggregated at the census tract level

In addition, AR/SRU maintains ongoing relationships with several Chicago-area research and policy organizations that have worked with some of the agencies the Project believes it will be necessary to contact. These relationships have proved helpful in learning about agencies and developing contacts within them. For example, AR/SRU staff communicate on a regular basis with the Chapin Hall Center for Children (a University of Chicago-based child welfare research group), the Consortium for Chicago School Research (a citywide policy group, also based at the University of Chicago, that is assessing the pace and impact of school reform in Chicago), and the Chicago Area Research Association (a monthly gathering of substance abuse and crime researchers in the Chicago area, sponsored by Treatment Alternatives for Special Clients, known as T.A.S.C.). AR/SRU staff also attend monthly meetings of technical groups consisting of city and county officials who are developing and refining criminal justice agency information systems. These researcher-to-researcher, researcher-to-practitioner, and policy-oriented links are an invaluable and ongoing component of AR/SRU's work. They keep project staff abreast of developments in the various agencies and promote familiarity of the Project and AR/SRU staff with the agencies and recordkeeping officials.

During AR/SRU's efforts to gain access to agency records, the unit learned how to develop a cooperative and mutually beneficial relationship with an agency. Without this relationship, data access would be difficult, if not impossible. AR/SRU created a seven-step process, which is illustrated in figure 4.1 and summarized below. Although this process implies an orderly progression in which each step builds upon the previous one, it is necessary to remain flexible should an unusual situation arise.

Figure 4.1



Step 1: Background Review

The first step toward obtaining data is to learn as much as possible about an agency through written sources and communication with others familiar with that agency. The knowledge gained will later facilitate dialogue with the agency and serve to establish the credibility of the researcher. In compiling background information, emphasis should be placed on learning both official characterizations and outside perspectives on the agency.

Step 2: Establish Agency Contact

Armed with detailed background information, the researcher is ready to make the first contact. Depending on the researcher's objectives and what has been learned about the agency, various approaches may be warranted. It is possible that enough background has already been learned about the agency to make a specific data request to a specific individual (this may be the case with agencies that have specialized and well-defined data collection missions, such as the U.S. Census Bureau). More likely, the researcher will need to learn more about the agency and whom to contact. In these instances, the purpose of the initial contact is to begin a dialog to discover the types of data available and how they may be relevant to the project. The researcher should also be prepared to share information about the research at hand with the agency and to lend expertise.

Step 3: Interview Knowledgeable Persons

Once the first contact has been made, the next step is to identify other agency personnel, if necessary, to understand the full scope of data maintained by the agency, the logistics of accessing the data, and procedures for obtaining specific permission to access the data. In larger, more complex agencies, key administrators, managers, and information specialists should all be contacted to obtain a broad perspective on data collection. It also is useful to speak with, and if possible observe, data input/output operators and record clerks. The people with hands-on, day-to-day data experience can provide valuable information about system capacities and data limitations. Their perspectives usually are more realistic than those of administrators and managers.

Step 4: Obtain System Documentation

Having gained the acceptance of agency personnel, the researcher will be in a position to move into a more sensitive or intrusive learning phase, beginning with documentation of data systems. Documentation may include data dictionaries (layout schema), data system users' manuals, input forms, and sample reports. Review of this material will help further define agency capabilities and data limitations and form the basis for further questions. This process will serve to define the scope of data collection in more concrete terms. It also will help the researcher understand how well the data system captures critical events and decisions made in the field and what types of information remain unrecorded.

Step 5: Access Test-Case Data

In much the same way that researchers pretest data collection instruments, AR/SRU has found it useful to access test-case data before attempting to obtain data on study participants. Such a hands-on trial period can familiarize project staff with data sources and is a good mechanism for assessing data quality. Test-case access also allows data collection problems to be worked out before actual data collection begins. Depending on the agency data systems, test-case data may be received in automated format or may involve project staff coding data from manual data sources.

Step 6: Automate Test-Case Data

Presuming the requester is able to obtain test-case data from manual and/or automated agency data sources, the next step is to translate the data into a format that meets the requester's software and data analysis specifications. If test-case data are not available, automating data information will have to be postponed until "live" data exist.

Step 7: Collect and Assess Data

When it is time to collect data on study participants, a clear plan should be written specifying the Project's request and the agency's responsibilities. In some instances, this may be written into the formal agreement between the parties. When the agencies are operating under less formal, good-faith agreements, the plan will be useful for reference and management purposes. This plan should include an enumeration of the specific tasks involved in data collection and a timetable for completion of tasks. All fees should be determined in advance (e.g., agency programming or per record search and production costs). If specified by the agency, the requesting organization should provide copies of release forms from study participants. As data are collected, quality assessment measures should be implemented to determine, to the extent possible, the accuracy (validity), consistency (reliability), and completeness of pertinent agency record data. Various analytic techniques can be used to investigate data quality (see table 4.2).

Table 4.2
Analytic Tools and Resources for Agency Record Data Quality Assessment

Agency Based	Outside the Agency	Researcher
Audits, technical reports, research reports	Users of agency data (researchers, officials from other agencies)	Research data-processing systems
Knowledgeable persons in the agency	Auditors of agency data	Analyze agency data (trends, frequencies, logical edit checks)
		Triangulation—compare with other data sources

The process for gaining access to agency records serves as a blueprint for staff activities. Although it is not always implemented in the regular sequence of steps that this discussion has implied, it keeps research activities focused and promotes efficiency. AR/SRU's documentation of this process is an important contribution to the Project's efforts to institutionalize project knowledge and history in order to promote an understanding of the Project's work and the development of a practical knowledge base for other researchers and research managers.

Reference

Coldren JR, Markovic JD, McElvain T, and Lazer M (1995). *A Systematic Process for Accessing and Assessing Agency Records for Longitudinal Research.* (Available through the Project on Human Development in Chicago Neighborhoods' Boston office; call 617-432-1102.)

Part Two: Measures Development

Chapter 5. A Community Survey Approach to Ecological Assessment: Results From a Pilot Study

by Jacqueline Barnes, Ph.D., Robert J. Sampson, Ph.D., Daniel Kindlon, Ph.D., and Albert J. Reiss, Jr., Ph.D.

There are strong theoretical reasons for believing that structural and cultural dimensions of community social organization are relevant to understanding the dynamic nature of substance abuse behavior, delinquency, and crime (Reiss, 1986; Sampson, 1994; Sampson and Groves, 1989). The degree of community organization/disorganization—a community's ability to realize the common value of its residents and maintain effective control over their behavior (Sampson and Groves, 1989; Sampson, 1992) as well as the degree of ecological segregation leading to differences in a community's cultural organization—helps shape expectations for conduct. This chapter describes a pilot test designed to assess the reliability and validity of a new instrument for ecological assessment. The broad dimensions defined by a theory of social disorganization were operationalized in a survey instrument administered to residents of Chicago neighborhoods by the Project. This instrument constituted a preliminary step in assessing the reliability, validity, and practical utility of direct survey measurement of multiple aspects of neighborhood social organization within the context of a major American city.

Method

Sample. The pilot survey was conducted in six Chicago census tracts—high-income white, Hispanic, and African-American neighborhoods, and low-income white, Hispanic, and African-American neighborhoods. Income information was obtained from the 1990 census and used to create a trichotomous measure (low, medium, and high) of socioeconomic status. A neighborhood had to have at least 75 percent of the race/ethnic group (e.g., more than 75 percent white) in question. Eight blocks from each of the six neighborhoods were randomly chosen, and interviewers called at every house in the random blocks until 25 community survey interviews were completed for each neighborhood type, with visits made at different times of day. Interviewers were limited, however, in that only four interviews could be completed in any one block, or one interview per apartment building. Block starting points were randomized. To limit cost, callbacks were not made, and all interviews completed were first attempts. The representativeness of the sample was not the main purpose of the study; rather, the utility of the interview, the nature of the scales, and the creation of the socioeconomic categories made it likely that the sample would have a range of economic/educational backgrounds.

The interviewers working in Hispanic neighborhoods were fluent in both English and Spanish. If any interviewer encountered an individual who did not speak either of these languages, the next house was visited. For this pilot phase, it was not possible to interview sufficient numbers from other cultural groups to permit separate analysis.

An interview was conducted with any knowledgeable adult (age 18 or older) who was a resident of the household. If an individual under age 18 answered the door, the person was asked to identify an adult within the household who would be willing to complete the interview. A respondent was paid $10—mailed following completion of the interview—for participation.

The total sample of 152 interviews was obtained from 402 eligible households (a 38-percent response rate). The most common reasons for not completing the interview, which had to be done on the spot, were refusals (102), requests to call back (53), or no adult at home (40).

Twenty-five interviews were completed in each of the six neighborhood tracts, with an additional two from the high socioeconomic Hispanic neighborhood. The majority of the interviews, 125, were completed in English, with 27 conducted in Spanish. Almost two-thirds of the respondents were female (96, or 65 percent), and their ages ranged from 18 to 83, distributed as follows:

- 18–25: 26, or 17 percent.
- 26–40: 52, or 34 percent.
- 41–55: 34, or 22 percent.
- 56–70: 29, or 19 percent.
- 71–83: 11, or 7 percent.

There was more than one adult living in the household for the majority of the sample (134, or 88 percent). Of these, most had at least two other adults (114, or 85 percent), and at least one other adult was at home in 104 (78 percent) of the 134 homes in which 2 or more adults lived.

Procedure. The instrument was designed to be administered as a face-to-face structured interview that took an average of 45 minutes. To reduce potential problems with literacy, the interview questions were read aloud to respondents, who also could follow along with their own copies, available in English and Spanish versions.

The Spanish version of the interview was the product of a translation, back translation, and committee review process designed to meet the needs of the predominant Spanish-speaking group in Chicago, who are of Mexican origin.

The majority of the items had scale responses, similar to Likert, that asked, for example, "How likely it is that neighbors would break up a fight?" (a four-point scale from "very unlikely" to "very likely") or "How much of a problem is public drinking?" (a four-point scale from "a big problem" to "not a problem"). All response scales of four or more points

were presented to respondents on cards so respondents could examine them while replying. The questions were scored so that a high score represented a better outcome when directionality could be assumed.

The interview questions were sequenced to provide some logic for respondents and were not organized according to theoretical constructs. Initial questions covered relationships in the neighborhood, then physical and observable neighborhood attributes, presence of institutions and membership of organizations, ideas about neighborhood residents' values, and, lastly, demographic aspects of the household.

Statistical analyses. The internal consistency of each predefined scale derived from the interview was examined by calculating Cronbach alphas (Cronbach, 1951). These same items were all entered concurrently into a principal components factor analysis with varimax rotation to study the scales that would emerge without any a priori expectations. Factors with an eigenvalue of 1 or more were retained, and all items that had a loading of 0.40 or greater were considered part of a scale. These analyses were initially conducted separately for the Spanish-speaking respondents and the English-speaking respondents, with similar results. Thus, the data reported in this chapter are based on the entire sample.

Results

Reliability. Of the 17 scales that had been created to represent theoretical constructs of neighborhood structural and cultural organization, 11 (65 percent) had a Cronbach alpha of 0.70 or greater, and 4 (23.5 percent) had an alpha between 0.60 and 0.69 (see table 5.1). These results demonstrate the reliability of these scales in terms of the consistency with which individuals answer conceptually related questions in a similar manner.

The scales with the greatest internal consistency were those describing Social Capital (alpha 0.89), one aspect of Social Cohesion, and Local Social Participation (alpha 0.85), included under Informal Social Control. Also in that dimension was Informal Monitoring/Surveillance (alpha 0.82). These scales included questions about people who are willing to help neighbors in tangible ways, such as lending a small amount of money, watching out for trouble, and watching children in the neighborhood (Social Capital); neighbors mixing socially and giving each other advice (Social Participation); and neighbors who become activists when antisocial behavior is noted, such as a fight breaking out, people making graffiti, or children truant from school (Informal Monitoring/Surveillance). Neighborhood services and the residential stability of the area also were described consistently (Health Resources and Services, 0.74; Recreational Services, 0.81; Residential Stability/Change, 0.79).

Reports of the antisocial aspects of the neighborhood were somewhat more variable. Reliable descriptions of Social Disorder were made in that Public Incivilities and Neighborhood Disorder–Crime were among the highest reliability coefficients (alpha 0.86 and 0.79). Questions included in these constructs asked residents to judge how much of a problem trash, public drinking, drug use or sales, and other incivilities were and to report the extent to

which crimes such as fights with weapons, gang disputes, rape, or mugging were in their neighborhood. However, questions pertaining to personal experiences of crime were not so consistently associated (Personal Victimization, alpha 0.43). Scales that included questions about antisocial values or behaviors of neighborhood residents had acceptable but only moderate reliability. For example, Antisocial Role Models (General), asking how many people in the neighborhood made money by illegal means such as fencing stolen goods or selling drugs, had a coefficient of 0.62; Alienation/Anomia had a coefficient of 0.68; and Social Cleavages/Segregation, including items such as the police not caring, lack of trust between neighbors, and groups being in conflict, had a coefficient of 0.67.

Conflict Subcultures was the only subscale close to zero. This may reflect the fact that it was a combination of attitude items and neighborhood descriptions.

Validity. In order to examine the validity of the item groupings without any predefined constructs, a factor analysis was completed, producing 23 factors, 15 of which contributed 2 percent or more of the variance. Table 5.2 shows the content of two major factors, including items loaded with values of 0.40 or greater.

The factors (see table 5.2) have been labeled to illustrate their relationship with the theoretically defined scales. Most of the items that were expected to group together did so, although the major factors reflected a combination of the a priori scales. Factor 1 (Social Participation/Cohesion), explaining 17 percent of the variance, represented the majority of Local Social Participation (alpha 0.85), Social Capital (alpha 0.89), and Informal Monitoring/Surveillance (alpha 0.82), with one item from Density of Acquaintanceship (alpha 0.62). Factor 2 (Neighborhood Disorder), with an eigenvalue of 7.4, contributed 9 percent to the variance. It consisted of items from three of the scales under the general heading of Social Disorder, Neighborhood Disorder (alpha 0.79), Public Incivilities (alpha 0.86), and Social Cleavages/Segregation (alpha 0.67), but not Personal Victimization. The latter (alpha 0.43) was reflected partly in a much smaller factor (factor 7, Street Safety).

The next three factors were strongly associated with predefined scales. Factor 3 (Services, eigenvalue 4.3) included all the items from Organizational/Political Structure–Health Resources and Services (alpha 0.74), while factor 4 (Recreation, eigenvalue 3.5) included nearly all the items from Organizational/Political Structure–Recreational Services (alpha 0.81). Factor 5 (Values About Teenagers, eigenvalue 2.9) included all the items from Normative Age-Graded Expectations (alpha 0.73), with the exception of the item about teens damaging property, which had little variance. The remainder of Public Incivilities formed factor 6 (eigenvalue 2.6, 3 percent of variance), and the Antisocial Role Models scale from Cultural Structure was replicated exactly in factor 8 (eigenvalue 2.1, 3 percent of variance). Alienation/Anomia from the Social Cohesion construct and Organizational Mobilization Capacity from the Organizational/Political Structure construct also were represented by factors 9 and 10, respectively.

Discussion

The internal consistency analyses suggest that the theoretical constructs, as they have been defined, lead to consistent patterns of responses from individuals in nearly all cases. This provides support for the theoretical reality of community/neighborhood characteristics and their relevance for residents. Questions that asked about ways in which neighbors helped one another, mixed socially with each other, and were likely to intervene in observable trouble all yielded reliable scales. Reports of services available in the local area and the extent to which there had been changes or stability in the neighborhood also showed substantial interrelationship in predicted ways. These results give substance to the theoretical constructs of Informal Social Control, Social Cohesion, and Residential Stability/Change. There were cohesive responses to questions asking residents to judge whether observable problems (Public Incivilities) and criminal activities (Neighborhood Disorder) gave substance to the Social Disorder construct. There may, however, be some problems when residents are asked to judge what the values of neighborhood residents are, or the extent to which they feel alienated from mainstream culture. Personal experience as a victim of crime does not appear to form a viable neighborhood construct and may have much more to do with individual or family variability than neighborhood characteristics.

In a few cases it was clear that an item did not contribute to a scale; these items have been removed from the final measure. For example, a question asking how the neighborhood compared to others in the city had been included with items about changes in levels of safety (Residential Stability/Change) but had a very low item-total correlation. Similarly, questions about neighbors not caring and going their own way, included under Social Capital with reversed scoring, did not contribute.

The factor analysis supports the concept that neighborhoods have a clear identity in terms both of positive aspects (Social Cohesion) and antisocial activities (Social Disorder). Nevertheless, within these two broad conceptualizations, the subdivisions that are theoretically driven may be reflecting neighborhood characteristics in ways that are more subtle than the reality. For instance, the strong loading of items in the Social Participation/Cohesion factor suggests that the theoretical distinction between Informal Social Control and Social Cohesion may, in real life, mean that neighbors either help each other or they do not. If there are people who are reliable and will help each other (Social Capital), then they also step in to prevent local trouble (Informal Monitoring/Surveillance) and mix socially (Informal Social Control–Local Social Participation).

Similarly, the second major factor included items that had been labeled Social Cleavages/Segregation (e.g., groups in conflict), Neighborhood Disorder–Crime (fights with weapons a problem), and Public Incivilities (drug sales a problem).

Overall, these results indicate that properties of neighborhood structural and cultural organization can be reliably reported using standard survey research techniques. Nevertheless, the present study is limited with respect to the small sample size and the low

response rate inherent in a design that did not incorporate callbacks, and because the assessments of reliability and validity were conducted at the individual level and not analyzed at the neighborhood-variation level. Because census data have severe limitations, these results are encouraging. Hence, future implementation of survey instruments designed to assess neighborhood organization seems promising, especially when the limitations of the present research can be overcome. An important goal is thus to design studies in which multidimensional properties of social organization can be studied at the neighborhood level of analysis.

References

Anderson E (1990). *Streetwise: Race, Class, and Change in an Urban Community*. Chicago: University of Chicago Press.

Buckner J (1988). The development of an instrument to measure neighborhood cohesion. *American Journal of Community Psychology*, 16, 771–791.

Bursik RJ (1988). Social disorganization and theories of crime and delinquency: Problems and prospects. *Criminology*, 26, 519–551.

Cronbach LJ (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 44, 373–393.

Furstenberg F (1993). How families manage risk and opportunity in dangerous neighborhoods. In: Wilson WJ (ed.), *Social and the Public Agenda*. Newbury Park, CA: Sage Publications, Inc., pp. 231–258.

Kornhauser R (1978). *Social Sources of Delinquency*. Chicago: University of Chicago Press.

Reiss AJ, Jr. (1986). Why are communities important in understanding crime? In: Reiss AJ, Jr., and Tonry M (eds.), *Communities and Crime*. Chicago: University of Chicago Press, pp. 1–34.

Sampson RJ (1994). The community. In: Wilson JQ and Petersilia J (eds.), *Crime*. San Francisco, CA: ICS Press, pp. 193–216.

Sampson RJ (1992). Family management and child development: Insights from social disorganization theory. In: McCord J (ed.), *Advances in Criminological Theory* (vol. 3). New Brunswick, NJ: Transaction, pp. 69–93.

Sampson RJ and Groves WB (1989). Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology*, 94, 774–802.

Sampson RJ and Lauritsen JL (1994). Violent victimization and offending: Individual, situational, and community-level risk factors. In: Reiss AJ, Jr., and Roth J (eds.), *Understanding and Preventing Violence: Social Influences* (vol. 3). Washington, D.C.: National Academy Press, pp. 1–114.

Shaw CR and McKay HD (1942). *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.

Skogan W (1986). Fear of crime and neighborhood change. In: Reiss AJ, Jr., and Tonry M (eds.), *Communities and Crime*. Chicago: University of Chicago Press, pp. 203–229.

Wilson WJ (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: The University of Chicago Press.

Table 5.1

Internal Consistency of A Priori Subscales Expressed as Cronbach's Alpha and Item/Total Pearson Correlations

1. Social, Economic, and Demographic Structure

(i) Residential stability/change — Alpha 0.79

0.66	neighborhood look changed
0.61	quality of people in neighborhood changed
0.61	personal safety changed
0.59	how will neighborhood be in a few years
0.53	level of police protection changed
0.24	how does neighborhood compare with others in the city

2. Organizational/Political Structure

(i) Health resources and services — *Alpha 0.74*

0.68	does neighborhood have family health service
0.56	does neighborhood have alcohol/drug program
0.52	does neighborhood have family planning clinic
0.50	does neighborhood have day care

(ii) Recreational services — *Alpha 0.81*

0.67	does neighborhood have recreational/afterschool programs
0.66	does neighborhood have afterschool program
0.62	does neighborhood have summer recreation program
0.52	does neighborhood have any good parks
0.48	teens can only hang out
0.48	does neighborhood have day-care services
0.45	no place to play but the street
0.33	does neighborhood have a community center

(iii) Organizational mobilization capacity — *Alpha 0.77*

0.65	would neighbors organize to fight cuts
0.65	would neighborhood organization be effective

3. Informal Social Control

(i) Local social participation — *Alpha 0.85*

0.75	how often favors exchanged
0.62	how often watch property
0.65	how often ask neighbors for advice
0.63	how often go to neighbors' parties
0.67	how often visit with neighbors

Table 5.1
continued

(ii) Informal monitoring/surveillance — *Alpha 0.82*

0.69	would neighbors stop children making graffiti
0.67	would neighbors report truants
0.62	would neighbors scold a disrespectful child
0.58	would neighbors break up fight

(iii) Density of acquaintanceship — *Alpha 0.62*

0.54	number of adults you recognize
0.51	number of children you recognize
0.40	number of friends in neighborhood
0.35	easy to spot strangers
0.24	number of families in neighborhood know each other
0.12	number of relatives in neighborhood

4. Social Cohesion

(i) Social capital — *Alpha 0.89*

0.76	are people willing to help
0.73	do neighbors watch out for kids
0.70	is this a close-knit neighborhood
0.69	do neighbors watch for trouble
0.67	could you borrow $30 from a neighbor
0.65	if there is a problem will neighbors deal with it
0.60	are there adults to look up to
0.56	can neighbors be trusted
0.49	people will take advantage
0.46	no one in neighborhood cares (reversed)
0.38	people go their own way (reversed)

(ii) Alienation/anomia — *Alpha 0.68*

0.51	only easy and hard ways to make money
0.50	laws made to be broken
0.44	anything is OK if no one hurt
0.40	live for today

5. Social Disorder

(i) Neighborhood disorder–crime — *Alpha 0.79*

0.67	fights with weapons
0.68	gang fights
0.57	violent arguments
0.53	rape
0.39	mugging

Table 5.1
continued

(ii) Public incivilities		*Alpha 0.86*
0.74	trash	
0.69	vacant buildings	
0.66	public drinking	
0.66	drug use/sales	
0.65	groups of teens/adults causing trouble	
0.52	graffiti	
(iii) Social cleavages/segregation		*Alpha 0.67*
0.58	police not caring	
0.42	groups in conflict	
0.42	police brutality	
0.41	lack of trust	
(iv) Personal victimization		*Alpha 0.43*
0.42	home broken into	
0.38	many fear going out at night	
0.35	risky to walk alone	
0.12	something stolen outside home	
0.07	family member victim of violence	
6. Cultural Structure		
(i) Normative age-graded expectations		*Alpha 0.73*
0.59	how wrong for teens to use alcohol	
0.54	how wrong for teens to use marijuana	
0.53	how wrong for teens to have sex	
0.51	how wrong for teens to skip school	
0.20	how wrong for teens to damage property	
(ii) Antisocial role models (drugs)		*Alpha 0.79*
0.66	number of people making money from fencing	
0.66	number of people making money selling drugs	
(iii) Antisocial role models (general)		*Alpha 0.62*
0.58	number of people making money from fencing	
0.56	number of people making money from drugs	
0.48	number of people making money placing bets	
0.05	number of people have regular job (reversed)	

Table 5.1
continued

(iv) Conflict subcultures *Alpha 0.07*

0.15	areas where trouble expected
0.06	fights between friends no one's business
0.02	gang fights should be ignored by police

Table 5.2

Results of Principal Component Factor Analysis of Neighborhood Survey with Varimax Rotation, including Factors with Eigenvalues Greater than 1 and Items with Loading of 0.40 or Greater

Factor 1: Social participation/cohesion

Eigenvalue 14.4 17% variance

0.73	a close-knit neighborhood
0.72	neighbors watch out for kids
0.72	could borrow $30 from a neighbor
0.72	how often favors exchanged
0.71	people are willing to help their neighbors
0.70	number of friends in the neighborhood
0.66	how often neighbors watch each other's property
0.63	adults to look up to
0.63	how often ask neighbors for advice
0.62	how often are neighborhood parties
0.61	if a problem neighbors will deal with it
0.60	how often visit with neighbors
0.59	neighbors watch for trouble
0.59	would neighbors report truants
0.57	would neighbors stop children making graffiti
0.52	neighbors can be trusted
0.51	would neighbors break up a fight
0.45	would neighborhood organization be effective
0.44	number of adults you recognize
0.44	would neighbors scold disrespectful children

Factor 2: Neighborhood disorder

Eigenvalue 7.4 9% variance

0.83	how often gang fights in neighborhood
0.72	how often fights with weapons in neighborhood
0.72	how often violent arguments in neighborhood
0.70	how big a neighborhood problem is groups in conflict
0.63	how big a neighborhood problem is teens/adults causing trouble
0.58	how big a neighborhood problem is lack of trust
0.51	how big a neighborhood problem is drug use/sales
0.45	how big a neighborhood problem is cops not caring
0.43	how big a neighborhood problem is graffiti
0.41	how often rape in neighborhood

Chapter 6. Systematic Social Observation of Chicago Neighborhoods

by Albert J. Reiss, Jr., Ph.D.

A principal research objective of the Project is to measure how growing up in different types of neighborhoods affects human development. This requires identifying local neighborhoods in Chicago and then measuring their properties that directly and indirectly affect human growth and development.

Block and census tract data from the 1990 Census of Population and Housing were the primary source of data used to identify neighborhoods and their spatial boundaries. These data identified 80 neighborhood clusters in the city of Chicago from which the subjects for the study were selected.

The design called for specifying the properties of neighborhoods by three independent methods of data collection and analysis: (1) block and census tract data from the 1990 census of population and housing, (2) a community survey of occupied dwelling units to select a sample of adults, and (3) systematic social observation of the neighborhood blocks. Below is an examination of the method of systematic social observation (SSO), followed by a comparison of the SSO profile with profiles derived from the other two methods of data collection on Chicago neighborhoods.

One may ask why it is necessary to use SSO as one of the multiple methods of data collection. The answer is inherent in the basic design of the study, which calls for examining neighborhood environment as well as individual and group effects on human development with annual observations for 8 years. A substantial number of urban families and individuals change their residence over 8 years; some will move to cities and towns lacking census tract information, and all who move outside the Chicago city limits will move to neighborhoods for which community survey data are lacking. SSO is a method to collect the information for movers at minimum cost.[1] Moreover, by having an SSO for all subjects, differences can be assessed between movers and stayers. With SSO measurement of all subjects, an SSO protocol can be completed each time there is home measurement.

SSO is a method of data collection by which trained observers using a standard protocol record the composition and behavior of a population or its sampled units.[2] The method was designed to collect data on natural social phenomena (Reiss, 1971). The basic model of SSO is analogous to that of sample social surveys (Reiss, 1968, 1975, 1976). Accordingly, SSO

---

[1]The cost of a community survey for each location to which a subject moves is estimated to exceed the cost of data collection for all subjects for the 8-year period.

[2]For a description of SSO, see "Systematic Observation of Natural Social Phenomena" in Herbert Costner (ed.), *Sociological Methodology*, San Francisco: Jossey-Bass, Inc., 1971, pp. 3-33.

specifies a sampling frame and a sampling plan, selects one or more samples, constructs and pretests a standardized protocol for data collection, trains observers as data collectors, creates a code for the information, and devises an analysis protocol. The basic difference between SSO and population censuses and sample surveys is that direct observation by a trained observer replaces respondent reporting to a survey interviewer (Reiss, 1968).[3] SSO differs from field research in that it systematizes the data collection for all units.

Several different techniques have been used to systematically collect observational data on neighborhoods and communities in Chicago. The first mode of data collection, the SSO, trained persons to code observations during or soon after the period of observation in a standardized schedule. This required an intensive period of training observers, using filmed or visual materials or actual field observation to achieve agreement with an experienced observer-rater (Reiss, 1971, 1979). The next innovation, by Taylor et al. (1984), characterized the residential site of each respondent by taking colored slides of the front and rear of the house; subsequently, two raters independently rated the physical features shown in the slides. Inter-rater reliability varied by item scored, but the ratings reached acceptable levels of reliability. Weisburd and Green (1994, 1995) conducted the first full-scale validation of their neighborhood areas using the observations and work experience of narcotics detectives to define small areas of drug sales (hot spots). The detectives' designation of drug-selling areas by type of drug was validated by data on arrests for possession or sales of different drugs. They then coded videotapes of these hot spots to characterize them.

A more structured method of observation was the Neighborhood Observation Schedule, developed and pretested to collect information on the residential and organizational characteristics for a population of face blocks. The observation schedule was designed by Jacqueline Barnes and Albert J. Reiss, Jr.,[4] and pretested in the Dorchester area of Boston, Massachusetts. A modified version of this schedule was prepared and administered under contract to the Project by the National Opinion Research Center (NORC) in the Chicago neighborhood sample.

The Dorchester survey was conducted in 1993. The basic unit of observation was the face block in five census tracts selected for the community survey. An observation schedule was completed by observers who began observation at one corner of a block and continued their observations to the next corner. Twenty-three Dorchester face blocks were independently

---

[3]One should bear in mind that interview surveys often ask for some observational data, e.g., type of housing and its condition.

[4]The schedule drew upon previous interview and observation schedules, especially a windshield survey developed by Ralph Taylor in Baltimore, Maryland, and the drug market survey in Jersey City, New Jersey, a joint undertaking of David Weisburd and Lorraine Green of Rutgers University and Capt. Frank Gajewski and Lt. Charles Bellucci of the Jersey City Police Department. Observation surveys in which the data are collected by observers in a drive-by auto are commonly designated "windshield surveys."

rated by two observers, and 52 blocks were rated by a single observer. The observer schedule collected information for 57 items making up 5 subscales:

- Street conditions (e.g., traffic).
- Residential housing characteristics.
- Neighborhood quality of life.
- Community vigilance.
- Commercial establishments and vacancies.

The main goals of the Dorchester survey were to establish inter-rater reliability for the face-block items and to investigate the ability of the SSO instrument to discriminate among blocks in a single inner-city community area. The reliability of each item was measured first in terms of the percentage agreement over the 23 blocks rated by two observers and, second, as a Kappa statistic. Cronbach's alpha also was calculated for all of the items in a subscale, regardless of the reliability or positive/negative assessment of any item.

Approximately half of the items (28 of 57) had a high proportion of identical responses. This was defined as any item in which at least 70 percent of the ratings were identical. An additional 14 items had moderate agreement: between 50 percent and 69 percent of the ratings were identical. For roughly one-quarter of the items (13 of 57), observers agreed less than one-half of the time. Agreement was higher for items that described the traffic volume, nature of housing or other buildings, presence of amenities, and presence of houses for sale. Agreement was less for the type and number of people present and the quality of the environment, e.g., noise level or sanitation.

Kappa statistics were calculated for the five subscales developed for their face validity. There was substantial variation in Kappa values in all of the subscales. This was in part due to the low number of positive responses to some items as well as to disagreement between the raters. Nevertheless, each of the subscales demonstrated some validity. The measure discriminated among face blocks in the small and relatively homogeneous community of Dorchester. Moreover, when the block ratings were regressed on the number of homicides that took place in 1990, traffic conditions and quality of life in the neighborhood were significant predictors of homicide rates.

The Chicago SSO followed a different procedure of collecting observations. The procedure called for videotaping both sides of a face block from a van in 80 neighborhood clusters with some 27,000 face blocks. Each videotape was accompanied by an observation log that was completed by two observers, each viewing one side of a block. To facilitate matching, all logs with the videotapes contained a sequential ID number, the name of the street of the face block, the cross streets that defined the face-block street segment, the direction the vehicle

was heading, the time, and the odometer reading. These appeared at the beginning and end of each face block. The sequential ID number, the time, and the odometer reading stated in the audio and written on the observation log had to match for coding the face block. Coders were trained to code the videotapes of these numbered face blocks, completing a coding form of 56 items for each. Coders were trained and supervised in the coding task using an SSO Coding Manual (June 1995). The items were divided into 35 objective items (e.g., Is a public telephone visible?) and 21 subjective items (e.g., Would you characterize the peer group as a gang?). Ten percent of all cluster face blocks were viewed by two coders to detect and correct coding errors. A discrepancy rate of 2 percent was considered acceptable for the objective items, while up to a 10-percent discrepancy rate was acceptable for subjective items. When the rate exceeded 2 percent or 10 percent, respectively, the coding supervisor reviewed the videotape footage for the face block, adjudicated the discrepancy, and updated the data base with the corrected information.

Reliability coefficients are available for only one of the items coded during observation and their coding from the videotapes in the NORC data set. In that case, the coding of land use into residential, business/institutional, mixed residential and institutional, and a residual category had 71.7-percent exact agreement. The discrepancies occurred primarily for codes that bridged a primary land-use category. The disagreement for residential codes occurred mainly for the mixed residential/business and mixed residential/institutional codes.

Comparisons of the SSO with census information and the resident community survey are currently under way.

References

Reiss AJ, Jr. (1968). Stuff and nonsense about social surveys and observation. In: Becker HS, Greer B, Riesman D, and Weiss RS (eds.), *Institutions and the Person*. Chicago: Aldine, pp. 351–367.

Reiss AJ, Jr. (1971). Systematic social observation of natural social phenomena. In: Costner H (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass, Inc., pp. 3–33.

Reiss AJ, Jr. (1975, 1976). Systematic observation surveys of natural social phenomena. In: Sinaiko AW and Broedling LA (eds.), *Perspectives on Attitude Assessment: Surveys and Their Alternatives*. Washington, D.C.: Smithsonian Institution (1975:132–150); reprinted 1976, Champaign, IL: Pendelton Publications, pp. 123–141.

Reiss AJ, Jr. (1979). Systematic social observation in police research. In: Knutsson J, Kuhlhorn E, and Reiss AJ, Jr. (eds.), *Policy and the Social Order: Contemporary Research Perspectives*. Stockholm: National Swedish Council of Crime Prevention, pp. 271–305.

Taylor RB, Gottfredson S, and Brower S (1984). Block crime and fear: Defensible space, local social ties, and territorial functioning. *Journal of Research in Crime and Delinquency*, 21, No. 4, pp. 303–331.

Weisburd D and Green L (1994). Defining the drug market: The case of the Jersey City DMAP system. In: MacKenzie DL and Uchida CD (eds.), *Drugs and Crime: Evaluating Public Policy Initiatives*. Newbury Park, CA: Sage Publications, Inc., pp. 61–76.

Weisburd D and Green L (1995). Policing drug hot spots: The Jersey City DMAP experiment. *Justice Quarterly*, December.

Chapter 7. The Translation Process: Expanding the Utility and Validity of the Project Protocol

by Lorrie Rickman Stone, Ph.D., and Chris Payne, Ed.M.

There are compelling reasons to translate research instruments when studying diverse populations. Researchers who take the time and effort to translate instruments increase the reliability and validity of their assessments of diverse populations and improve completion rates by demonstrating an interest in and sensitivity to other cultures. The Project's decision to translate research instruments and field documents into Spanish came from early reviews of Chicago's population statistics (Census Bureau, 1992) describing 15 percent of the city's population as Spanish speaking, 4 percent of whom were linguistically isolated. It was only after the study began and there appeared to be larger numbers of Polish-speaking residents in Chicago than reported in the 1990 census that a decision was made to translate the instruments into Polish as well. This chapter describes the process used by the Project to translate the assessment instruments into both Spanish and Polish and offers impressions garnered from this experience.

Traditional Translation Methods

The literature on translation methods describes three primary techniques: (1) one-way translation, (2) translation by committee, and (3) double or "back" translation. The one-way translation technique usually involves the use of one bilingual individual to translate the original version of a document into the target language. The translator depends on knowledge of the language and information available through dictionaries and other references. Although this technique is simple and economical, it has several drawbacks. First, it relies exclusively on the interpretations of one individual. In a worst-case scenario, a translator lacking knowledge or understanding of a particular culture's semantics and idioms could produce a highly literal translation lacking the connotative meaning expressed in the original document. Further, Berkanovic (1980) found that Spanish versions of instruments translated using this method had lower reliability and lower bivariate correlation scores than the English version. Given these reasons, the one-way translation method is not preferred.

Translation by committee involves the use of two or more individuals familiar with both languages who separately translate the instruments into the target language. Translators may then meet to discuss differences in translated versions and arrive at a consensus. Another option is to permit an independent observer/expert to choose the version that seems most appropriate. Although this technique is more rigorous than the one-way translation, there may still be difficulty associated with translators *sharing* a view of the world based on similar backgrounds, education, and experience. This can be particularly problematic if the translators' worldviews and experiences are disparate from the experiences and background of the population for which the instrument was designed. Another limitation of this method may involve a reluctance on the part of committee members to discuss critically the work of

their colleagues. Having clear expectations for the work of the committee and acknowledging the potential difficulty this process may involve seem to facilitate completion of work via this process.

The preferred translation technique is known as the double, or "back," translation method. As with the committee method, multiple translators are required. In its simplest version, one bilingual translator takes the original version and translates it into the target language. Once completed, an independent translator takes the translated version and produces a new version of the original instrument in the original language of that instrument. Thus, at the end of this process, there are three versions of each instrument: the original version, the translated version, and the back translation. The researcher then compares the original version with the back translation to examine inconsistencies, mistranslated sentences, vague or lost meanings, or culturally inappropriate or insensitive wording. If differences are found, the researcher can consult with both translators to reach a consensus on the best alternatives.

Researchers may elect to do more than one iteration of both the original translation and the back translation. Research has shown that producing multiple iterations of the instrument using this translation technique takes the instrument through a number of filters and produces better overall versions of the instrument (Marin and Marin, 1991). However, using multiple translators is no guarantee that the translation to the target language conveys the appropriate meaning. If translators share a common set of linguistic rules, they may be apt to miss important language/cultural nuances requisite for appropriate translation. Marin and Marin (1991) report this problem more frequently when the translators are bilingual but not bicultural. Another common pitfall associated with this technique is the proclivity for translators to attempt literal translations that fail to represent the desired connotative meanings. Often, translations done literally result in language that is stilted and awkward at best.

The Project's Translation Techniques

In translating research instruments and field documents into Spanish and Polish, the Project combined two of the basic techniques, translation by committee and back translation, enhancing the methodology of each technique to improve the quality of the translation and facilitate management of the translation process. First, research assistants fluent in the target languages were solicited to become part of the translating team. Then professional translators were contracted to conduct the back translation; they were selected on the basis of their relevant experience, knowledge of and exposure to Spanish and Polish culture, and willingness to collaborate with Project staff.

Translation by Committee

Although the procedures for translating documents were essentially identical for both languages, the experience of producing translations in each language was quite different and is described in a subsequent section. A total of 22 measures (37 distinct measures, which

include different versions of the same measure) and several field documents, totaling approximately 325 pages, were prepared for translation. Eleven measures had different versions for different ages or respondent types. Versions of the same measure were grouped together, and identical sections/wording were highlighted for the translator to facilitate consistency of the translation across versions.

English versions of instruments were distributed to bilingual research assistants to produce the translations. They were instructed not to collaborate with one another during the initial translation process, but they were able to use dictionaries, books, magazines, and other professional texts.

Once the initial translation was completed, a committee was formed to critically review and edit the documents. Forming the committee of Spanish-speaking translators was particularly challenging because it was important for members of this task force to represent the variety of Spanish-speaking cultures living in Chicago. For example, population statistics indicate that while the largest portion of the Spanish-speaking population is Mexican, there is also a large contingent of residents from Central American countries, including Guatemala and Nicaragua, and from Puerto Rico, Cuba, and South America. The Project was fortunate to have research assistants representing most of these cultures on the task force. In addition, most task force members were both bilingual and bicultural.

Forming the Polish-speaking task force was significantly less complicated, primarily because the Polish culture is more homogenous than the Spanish-speaking culture. The Polish-speaking task force also included translators who were both bilingual and bicultural.

The Spanish and Polish committees were responsible for comparing differences in their respective target translations, discussing these differences, and, whenever possible, arriving at a consensus on the final translation. When differences could not be resolved, a senior researcher and/or professional translator was consulted to assist in the decisionmaking process. Generally, it was necessary to consult with one of these parties to clarify the intent of the question in the English version, to rephrase the question in the target language to reflect the conversational style in that language, and/or to agree on the various options for descriptors and idioms that have different meanings in different cultures. For example, questions posed in English and in a manner typical of conversational style in the U.S. may be offensive or, at best, culturally insensitive if posed in the same manner in Spanish. Further, one word may have several meanings in Spanish, depending on the cultural context. In these instances, it was important to determine which word best represented the concept to be conveyed and would be commonly understood (e.g., as used in "broadcast Spanish"), or whether it was necessary to list the various options for the word in the Spanish language. Similarly, some English words have no comparable word in the Polish language. When this occurred, it was necessary to agree on descriptive words or phrases that adequately conveyed the concept. The final documents represented the best efforts of the committee, the professional translator, and a senior researcher.

Back Translation

Back translations were conducted by professional translators. Translated versions of the measures were grouped, and versions of the same measure were given to the translator concurrently. The translator was not permitted to see the English version of the measure and was instructed to translate each instrument and rate the adequacy of each question using the following rating system: A = no improvement needed; B = may need improvement; and C = should be improved or clarified. Items in the measures were reviewed on the basis of their clarity, grammar, syntax, and cultural appropriateness/sensitivity. For any item graded B or C, the translator offered commentary and suggestions for improvement. This technique clearly detailed concerns regarding each question and facilitated discussions and final resolutions. Because of the volume of work to be done, the back translations were done onsite to provide the translator with easy access to the senior research team. Translations were dictated and transcribed at the end of each day. On subsequent days, the translator reviewed the transcription and proofed the final back translation and commentary.

After the back translation was completed, any differences between the English version and the back translation were highlighted, as were items coded as B or C. The translation committee then met to discuss and ultimately arrive at a consensus on the necessary changes. A translation team member was then assigned the task of creating a master copy of each measure in the target language and incorporating the recommended changes. Final copies of the measure were drafted and reviewed again during a final editing process.

Impressions

The research team, including the professional translators, viewed the exercise of translating the measures as difficult, but important, and rewarding. Translating the high volume of research protocols under time constraints was a more arduous task for the translation team than expected. More lead time in this effort could have prevented some of the fatigue experienced by the team. However, the Project benefited from using trained bilingual/bicultural research staff to translate research documents; their understanding of the measures, field experience, and general Project knowledge facilitated the translation effort. They readily comprehended the intent of most questions, could anticipate respondents' reactions to questions, and were able to work independently and with cultural sensitivity.

Another vital component in the translation process concerned cultural issues. In addition to focusing on differences in grammar, style, and language, the translators also interpreted and analyzed each item for cultural compatibility. To neglect cultural issues would be to impede communication between the interviewer and the respondent and could result in misunderstandings and data of questionable validity. For example, a child growing up in Chicago might associate certain Polish or Spanish words with the English equivalent, yet, to a person who just arrived from Poland or Latin America, the words would not make sense. Likewise, different dialects within Chicago communities could affect an individual's

understanding of the questions' intent. Cultural sensitivity was of utmost importance to the translators during this process.

One of the issues that required cultural sensitivity was use of the word "American." As Roberto Mendoza, a professional Spanish translator stated, "Latin American individuals have a hard time digesting the term 'American'.... To ascribe the term 'American' only to those who are native citizens of the United States of America could be misinterpreted as discriminatory." The translator was concerned that questions regarding "American customs" could insinuate "white customs" and therefore insult many respondents. Care was taken to ensure that questions regarding ethnicity were translated appropriately into Spanish and Polish.

Measures asking for information on educational background also required cultural consideration. For instance, "high school" in the U.S. is usually composed of grades nine through twelve and is equivalent to "secondary school." In Latin America, "secondary school" is only up to the ninth year. Many countries do not have "colleges," but they do have "universities." And how do you translate the American GED in Spanish and Polish? The classifications of preschool, kindergarten, and day care also present difficulties in translation. One way to solve the translation dilemma is to ask the respondent for the number of years in school instead of the grade or year.

Translators were also concerned about the English word "relationship," particularly in the Spanish translation. Translators stressed that the questions had to distinguish which kind of "relationship" the English word meant. Did the item refer to a romantic relationship? Or a love relationship without living together? A husband-and-wife relationship? Or was it a blood/family relationship? In the Spanish translation, the translator had to be sure the correct Spanish words were used for the one English word "relationship." Additionally, the English word "spouse" had to be adjusted for translation. In Latin America, "spouses" could refer to individuals who live together and are not married.

Measures focusing on violence and substance use also required cultural sensitivity. Many English words for drugs have no equivalent in Polish. The phrase "to be high on drugs" required extra attention in the Spanish version. In the first translation, the intent of the item was lost because the Spanish think of "being high" as floating, skydiving, or flying, not being under the influence of drugs. Other concerns surfaced when translating measures about abusive behavior. Abuse in the United States might not be considered abuse in another culture, depending on different styles of punishment and cultural values. In addition, words such as "hurt," "choke," "crime," and "probation" had to be translated with consideration given to item intent and dialect.

The decision to translate a research protocol and the methodology employed are critical decision points in research with culturally diverse populations. In the Project's experience, a modification of two traditional methods of translating instruments, translation by committee and back translation, proved to be the most efficient and effective means of executing this

task. The Project made a substantial investment in this process, and the benefits reaped have been noteworthy and often unexpected. By engaging in this process, the Project gained the confidence and respect of its bicultural staff members and many of its sample members and increased its integrity in the community. Finally, and of equal import, the process has sensitized the staff to cultural differences and to the principles of linguistics that govern communications.

Acknowledgments

On behalf of the Project, the authors gratefully acknowledge the contributions of the translation staff, whose dedication and commitment to this process was invaluable.

Professional Spanish Translator
Roberto Mendoza
Spanish-Speaking Task Force
Ana Marrero, Facilitator
Gladys Arana-Nelson
Raymond Aviles
Shawn Early
Kim Ferullo
Pamela Jones
Gillian Kane
Nelly Lopez
Ingrid Magan-Ing
Isaias Mercado
Nancy Perez
Zoyla Ramirez
Ana-Maria Tirado

Professional Polish Translator
Gregory Pekala
Polish-Speaking Task Force
Margaret Wieczorek, Facilitator
Agnieszka Kaplan
Jolanta Swieton

References

Berkanovic E (1980). The effects of inadequate language translations in Hispanics' responses to health surveys. *American Journal of Public Health*, 70, 1273–1276.

Brislin RW (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1, 185–216.

Bureau of the Census (1992). Census of Population and Housing, 1990: Summary Tape File 3. Washington, D.C.: The Bureau of the Census.

Campbell D, Brislin R, Stewart V, and Werner O (1970). Back-translation and other translation techniques in cross-cultural research. Paper submitted to the *International Journal of Psychology*.

Marin G and Marin B (1991). *Research with Hispanic Populations*, Newbury Park: Sage Publications, Inc.

Schoua AS (1985). An English/Spanish test of decentering for the translation of questionnaires. (Doctoral) *Dissertation Abstracts International*, 46, DA8511857.

Chapter 8. Infant Assessment

by Daniel Kindlon, Ph.D.

In October 1994, plans were established to expand the research protocol to include the youngest Project cohort. With assistance from Daniel Kindlon, initial literature investigations and consultations resulted in the selection of domains for study and investigators to assist in choosing relevant measures. The centerpiece of this study is a visual recognition memory task employed as a measure of the speed of information processing at age 6 months. This study also includes measures of infant health, health, temperament, the primary caretaker's (mother's, father's, or other family member's) health and rearing style, and observations about the home environment. Four senior infant researchers who attended a workshop in October 1994—Susan Rose, Jerry Kagan, Nancy Snidman, and Mary Rothbart—agreed to become advisers to this study. Ann Strohn, a Ph.D.-level consultant, a unit manager, and five research assistants were hired for the Infant Assessment Unit (IAU). This chapter describes the process of developing the infant assessment segment of the Project.

Background and Significance

There is considerable evidence that, for at least a subset of persons, antisocial behavior in adolescence and adulthood can be traced to determinants manifested in early childhood (Buka and Earls, 1993; Robins, 1978; White, Moffitt, Earls, Robins, and Silva, 1990). For preventative efforts to be maximally successful, it is important to establish the earliest age at which these determinants become predictive (Earls, Cairns, and Mercy, 1993; Earls and Jung, 1987; Loeber and Dishion, 1983). This study examines this question by conducting an intensive prospective study of an infancy cohort embedded within the Project.

The Project incorporates a multilevel research design examining characteristics of the community, school, individual, and family. Because the community-level variables affecting members of the infant cohort and their caretakers are being assessed independently as part of the overall Project design, additional measures derive from the individual and family level. Although analysis at the school level is less relevant for the infant cohort than for older participants, information about day-care placement will be obtained.

Thus, the infancy cohort will have data from all four levels, and, as such, these data can be analyzed in a manner analogous to the types of analyses proposed for antisocial outcomes in the older cohorts. This will include conventional epidemiological analyses as well as hierarchical linear modeling (Raudenbush and Bryk, 1992). The longitudinal component of such analyses will evaluate developmental trajectories for infants nested with families residing within neighborhood types or clusters.

Literature on early determinants has established three broad domains of individual factors related to later conduct problems, delinquency, or criminality (Buka and Earls, 1993; Kagan, 1991). These domains pertain to cognitive abilities, physical health/prenatal or perinatal

complications, and temperament. Measures of cognitive functioning in infancy, in addition to being predictors of outcome, have also been shown to be outcome markers of neurological insult or environmental deprivation (Jacobson, Jacobson, Sokol, Martier, and Ager, 1993; Rose, 1983, 1994). The important divisions among family-level variables are maternal social support/father involvement, maternal depression, and parental responsivity/disciplinary tactics (Buka and Earls, 1993). After nearly a year of literature investigation and consultation with some of the leading authorities in infancy research, beginning with the Project Infancy Measures Workshop convened in October 1994, a set of techniques for the measurement of the earliest manifestations of important constructs within these domains has been identified.

In the design of this component of the Project, two objectives have been identified. The first purpose is to examine the effects of prenatal and early postnatal risk conditions on health and cognitive functioning in the first year of life. Within the context of the multilevel design, this analysis will provide a measure of the infant's health and cognitive status in circumstances of varying risk. It is already known that infant mortality and low birth rates vary according to some of the same socioeconomic conditions associated with variations in delinquency and crime. The questions the Project seeks to answer concern the magnitude of factors operating at different organizational levels of influence—from community to family to parent—on infant development and an improved understanding of the mechanisms that may be established in the first year of life. These mechanisms that represent the first steps on developmental trajectories consist of behavioral characteristics (poor impulse control, low verbal skill, aggressive acts) and detrimental circumstances (school underachievement, peer rejection, drug use) that result in delinquent and criminal behavior (Loeber and Leblanc, 1990). The second purpose is to establish the links between early developmental processes and the onset of antisocial behavior in the preschool period or in the earliest years of regular school and to measure the strength of this developmental pathway.

Much of the empirical basis guiding the Project's emphasis on infant development is reviewed in Moffitt's (1993) efforts to distinguish early-onset antisocial behavioral profiles from those that typically arise during late childhood or early adolescence. In this taxonomy, which has broad support, a minority of teenage delinquents have a history of aggressive behavior that extends back to early childhood. Typically, the proportion of males with such a history constitutes less than 10 percent of a representative population. Yet their involvement in crime follows a distinctly different pattern from the majority with a later onset. The early onset group are more likely to commit offenses more frequently and to engage in deviant lifestyles for longer periods than the "adolescence-limited" youths in Moffitt's taxonomy. Thus, much of the significance for the study of infancy is based on being able to observe patterns of continuity and discontinuity and to isolate mechanisms accounting for early-onset deviant behavior before a propensity for delinquency and criminality is established.

It should be clear that although this longitudinal study does not in itself offer an intervention, variables for measurement are designed and selected with plausible and acceptable intervention strategies in mind. For example, selection of cognitive variables, as discussed below, is based on the knowledge that early programs of health and educational promotion,

such as Head Start and Project Begin, can offset the attrition in skill and capability that accompany abusive or ineffective parenting strategies and disadvantaged social and economic community conditions (Earls and Carlson, 1995; Earls and Barnes McGuire, in press).

Cognition. Many studies find that low IQ, especially low verbal ability, relates to antisocial outcome (Busch, Zagar, Hughes, Arbit, and Bussell, 1990; Denno, 1990; Glueck and Glueck, 1934; Moffitt, 1990; Reiss and Rhodes, 1961). Poor educational attainment, especially reading problems, also is associated with later delinquency (White et al., 1990). These associations appear to hold even after controlling for family factors (Moffitt, Gabrielli, and Mednick, 1988; Reiss and Rhodes, 1961). Work in infant cognition (see review by McCall and Carriger, 1993) indicates that variables assessed with the visual recognition memory (VRM) paradigm afford the opportunity to measure an aspect of cognition in infancy that predicts IQ in childhood and early adolescence (Rose, Feldman, and Wallace, 1992; Rose and Feldman, 1995). For example, in the Rose and Feldman study (1995), VRM scores for 39 infants at 7 months old showed a correlation of .45 with their full-scale IQ at age 11 years. Moreover, the paired-comparison procedure employed in the VRM paradigm appears to be a more sensitive and valid measure than other similar infant cognition measures, such as habituation paradigms employing only a single stimulus (Kagan, 1989; Laucht, Esser, and Schmidt, 1994).

This measure will be important not only as a predictor but also to the Project's first research objective through its use as an outcome measure of early/prenatal environment effects, such as prenatal brain insult or severe environmental deprivation. Research indicates that the VRM paradigm is sensitive to the effects of prenatal alcohol exposure (Jacobson et al., 1993), poor nutritional status (Rose, 1994) and premature birth (Rose, 1983).

VRM studies, whether predictive or outcome, have tended to involve relatively small samples and as such lack the statistical power for multivariate statistical analyses. Furthermore, they assess a constrained set of parental and social factors. In the proposed study, however, the Project will be able to build on the important findings from previous research to derive an improved understanding of the earliest occurring mechanisms that potentiate delinquent and criminal behavior.

Physical health. Prenatal and perinatal insult has been shown in several studies to result in the types of neurological impairment—mental retardation, attention deficit disorder, and learning disabilities—that are associated with later antisocial and delinquent behavior (Denno, 1990; Hack, Breslau, Weissman, and Aram, 1991; Klein, Hack, and Breslau, 1989), and some, but not all, longitudinal studies find an association between perinatal or birth complications and violent crime in adulthood (Buka, Tsuang, and Lipsett, 1993; Raine, Brennan, and Mednick, 1994). These studies have examined a limited and disparate set of birth/perinatal complications. As part of the health assessment, infants' medical records will be obtained to evaluate the presence of specific birth and pregnancy complications, including those identified in the research cited above as well as those used in calculating the Sheffield birth score (Carpenter, Gardner, Jepson et al., 1983), an important public health risk index

used in early intervention trials (Myerberg, Carpenter, Myerberg, Britton, Bailey, and Fink, 1995).

Higher prevalence prenatal risk factors include alcohol, tobacco, and drug use and low maternal weight gain during pregnancy (Institute of Medicine, 1990; Kleinman and Madans, 1985; National Institute on Alcohol Abuse and Alcoholism, 1990; Tronick and Beeghly, 1991). Where assessed, these behaviors are underreported on birth certificates (Serdula, Williamson, Kendrick et al., 1991) and will be obtained through interviews. Recent prevalence estimates for some of these factors find that 17 percent of mothers were reported to have smoked during pregnancy, approximately 20 percent used alcohol, and nearly 10 percent had a weight gain of less than 16 pounds, a level associated with poor outcome (Ventura, Martin, Taffel, Matthews, and Clarke, 1994).

Important postnatal risk factors to be assessed include lead ingestion (Baghurst, McMichael, Wigg et al., 1992), slow growth as an index of poor nutritional status (Rose, 1994), and head injury (Denno, 1990). Health status during infancy can be assessed through a combination of medical and health record data, maternal report of infant illnesses and injuries, and health-related quality of life (Behavioral Risk Factor Surveillance Survey, 1994).

Temperament. Preliminary investigations indicate that from among the many temperamental designations that exist, there are three constructs present in infancy that are relevant to the subsequent development of delinquent and antisocial behavior. These are:

- Fearfulness/behavioral inhibition.
- Approach/behavioral activation.
- Anger/distress to limits.

In proper situations, observed motivated behavior is the product of the balance between two systems. In the context of Gray's motivational model (1975, 1982, 1987), there are two limbic arousal systems, a behavioral activation system (BAS), and a behavioral inhibition system (BIS). BAS is sensitive to environmental cues of impending reward or relief of punishment. It is an approach system in the sense that it activates and energizes behavior (either approach or active avoidance). In contrast, BIS involves sensitivity to cues of impending punishment and frustrative nonreward. The initial effect of BIS activation is the interruption of ongoing BAS-mediated, goal-directed behavior. Quay (1993) has suggested that aggressive individuals are characterized by low levels of BIS activity and a relative dominance of BAS.

Work by Jerome Kagan and his colleagues (Kagan, 1991; Kagan and Snidman, 1991) shows that aspects of the fear/behavioral inhibition continuum can be assessed by looking at the infant's reaction to stimulation. Kagan's view is that before 6 months, cortical inhibitory

systems are not operating and one gets a window on the arousability of the limbic system by observing motor output in response to stimuli. He and his colleagues have developed reactivity paradigms that code infant behavior in response to stimuli such as mobiles of increasing complexity. Children who show high levels of positive arousal are those who later will show low levels of anxiety/behavioral inhibition. It is hypothesized that if these children are raised in environments with high levels of crime and drugs, their higher thresholds for becoming anxious place them at greater risk for displaying antisocial behavior (Kagan, 1991). Work by Kochanska (1993) also suggests that individual differences in these types of regulatory traits play a role in the development of "socialization relevant" traits such as empathy and guilt.

In a similar vein, Rothbart, Ahadi, and Hershey (1994) have used a technique involving latency to grasp at an interesting toy, thought to reflect individual infantile differences in the capacity to plan and to suppress inappropriate approach responses, as a measure of temperamental regulatory ability. They find that shorter latencies to grasp are associated with lower levels of empathy at later ages.

Another construct assesses individual differences, thought to be temperamental in origin, in the extent to which anger or distress to frustration or limitations is manifest. The research on this domain, although more equivocal than in those previously discussed, suggests that this is a risk factor for later antisocial behavior and may be measurable in infancy (Caspi, Henry, McGee, and Moffitt, 1995; Gunnar, Porter, Wolf, Rigatuso, and Larson, 1995; Rende, 1993; Rothbart et al., 1994).

Parental responsivity and disciplinary tactics. Among family-level measures, the most powerful predictor of later delinquency and antisocial behavior is parenting practices, including poor discipline and parental rejection (see review by Buka and Earls, 1993). Higher severity of parental punishment, for example, has been shown to be related to higher levels of aggression toward kindergarten peers (Strassberg, Dodge, Petit, and Bates, 1994). The degree to which the parent uses verbal or physical aggression as opposed to verbal reasoning to control the child's behavior can be assessed by an interview (Straus, 1979). Emotional and verbal responsivity of the mother can be assessed by systematic observation (Bradley and Caldwell, 1980). In addition, sessions in the home will be videotaped, and the interactions observed in the home can be coded at a later time.

Maternal social support/father involvement. The degree to which the infant's primary caretaker receives social support, such as help with daily tasks, emergency child care, and satisfactory relationship with a supportive person, is associated with increased health of both mother and child (Pascoe, Chessare, Baugh et al., 1987; Pascoe and Earp, 1984). Previous work has shown that for African-American children, a high level of father involvement is associated with increases in IQ, even when family income is statistically controlled, and that this construct may be assessed through maternal report (Yogman, Kindlon, and Earls, 1995). It is conceivable that the high levels of stimulation typically involved in father-infant play facilitate cognitive development, which has been reported for vigorous tactile stimulation in

preweaning rats (Meaney, Aitken, van Berkel, Bhatnagar, and Sapolsky, 1988). Conversely, the effects of father involvement also may operate through increased support to the mother and facilitate effects mediated by her.

Maternal depression. Maternal depression has been associated with a lack of reciprocity to child cues, neglect, overcontrolled disciplinary practices, poorer attachment, and subsequent child psychopathology (Blehar, Lieberman, and Ainsworth, 1977; Downdy, Mrazek, Quinton, and Rutter, 1984; Welner and Rice, 1988). This can be assessed in a maternal interview that will allow distinctions to be made between transient postpartum depression and more chronic conditions.

Design and Methods

The infancy assessment, embedded as it is within the Project, will incorporate multilevel analytic features. As noted above, the two primary analytic aims are to conduct an outcome study and a prospective longitudinal study. The first of these examines cognitive outcome as a function of the effects of prenatal and early postnatal risk conditions. The second looks at the links between early developmental processes and the actual onset of antisocial behavior in the preschool period or in the earliest years of regular school and measures the strength of this developmental pathway. These latter analyses are explored through the use of hierarchical linear modeling (HLM; Raudenbush and Bryk, 1992), and the infancy study will add an important component by providing true prospective data beginning from birth. In HLM, a developmental path is associated with each person with respect to delinquent or conventional behavior; these paths vary over the population of persons, and variation in the directions and shapes of these paths depend on personal characteristics and social context. The infancy study will allow these trajectories to be analyzed from a true-age zero-time point.

Statistical Analysis of Power

This section describes needed sample sizes for the infant study under various assumptions about effect size, the distribution of the risk factors, and the correlations between risk factors. The first part considers individual-level risk factors; the second part considers community risk factors. The analysis takes as its illustrative case comparisons between risk groups on cognitive outcome in infancy as measured by a score on the VRM paradigm, which conforms to a continuous IQ-type scale.

Individual risk factors. The illustrative model examines the case where two risk factors, such as maternal alcohol use during pregnancy and poor nutritional status, predict VRM. Other covariates (e.g., parental education) may be needed in such a model, but these could either increase or diminish power. The Project will operate on the assumption that the covariates have no effect on power.

A further assumption is that for each risk factor, there are two groups: a low-risk and a high-risk group. Calculations for the effect size are made in standardized units: an effect size of .25 is one quarter of a standard deviation of the outcome variable. An effect size of .25 will be considered as small, .33 as moderate, and a .50 standard deviation as large.

Three possible distributions of risk are considered: equal distribution, p = .50; moderately skewed distribution, p = .25; and highly skewed distribution, p = .125. For simplicity, both risk factors are assumed to have the same distribution.

Two possible degrees of correlation between risk factors are considered: no correlation (a relative odds of 1.0) and moderately high correlation (a relative odds of 4.0, meaning that being at risk on factor 1 increases the odds fourfold of also being at risk on factor 2).

See table 8.1 below for the required sample sizes.

Table 8.1
Required Sample Sizes

Effect Size					
Small		Medium		Large	
Uncorrelated Risk	Correlated Risk	Uncorrelated Risk	Correlated Risk	Uncorrelated Risk	Correlated Risk
p = .50					
697	784	400	450	175	196
p = .25					
929	1,012	531	581	233	253
p = .125					
1,593	1,692	915	971	399	423

Relatively large effect sizes have been reported in the literature using VRM as an outcome (Jacobson et al., 1993; Rose, 1983, 1994). Nevertheless, it is desired that this study be able to detect a medium effect size with correlated risk factors that are moderately skewed. Such an analysis requires nearly 600 subjects.

Neighborhood risk. The power analysis for neighborhood risk uses as the standard error S ~ = ~ sqrt{{4 Delta}over J} with Delta ~ = ~ tau ~ + ~ sigma sup 2 over n, where J is the number of neighborhoods, n is the number of infants per neighborhood, τ is the variance between neighborhoods, and σ^2 is the variance within neighborhoods and is set to unity as above. It is assumed, based on the Project's ongoing work in Chicago, that n = 12 infants will be available per neighborhood.

The intracluster correlation—that is, the proportion of variance that lies between neighborhoods—will have a bearing on the number of needed neighborhoods. The intracluster correlation is given by rho ~ = ~ tau over{tau ~ + ~ sigma sup 2} ~ = ~ tau over{tau ~ + ~ 1}.

Three such correlations are considered: small (.02), medium (.05), and large (.10). It is predicted that the intracluster correlation (ICC) will be small, especially after controlling for relevant neighborhood-level covariates (e.g., poverty concentration) and that there will be small to medium effects of neighborhood risk factors.

The results are given below in table 8.2. Entries in the table are the required number of neighborhoods for power of .90. Total sample sizes of infants are in parentheses. The results assume that two types of neighborhoods are under comparison with equal numbers of neighborhoods in each group.

Table 8.2
Number of Neighborhoods (Infants) Needed for Analyses

	Effect Size	
	Small	Medium
Intracluster Correlation (ICC)		
Small	72 (864)	41 (492)
Medium	93 (1,116)	53 (636)
Large	121 (1,452)	68 (816)

The maximum number of available neighborhoods is 80. Given the small expected ICC, 80 will be sufficient to achieve high power, even with the small expected effects. Thus, from table 8.2 it can be seen that the total number of subjects needed will be 864 for small effect sizes and 636 for medium effect sizes.

Analytic Framework

Data analyses for the infancy study are for the most part identical to those used in the Project. The multilevel HLM analysis will be maintained along with conventional epidemiological analyses. Where appropriate, ordinal-level scale scores from interviews will be converted to interval-level data by Rasch modeling (Rasch, 1980; Wright and Masters, 1982).

There will be some forms of data in the infancy study that are not contained in the present Project design. Growth data, for example, will have to be converted to percentiles based on Centers for Disease Control and Prevention norms. Behavior during the mobile task will have to be coded from videotapes and inter-rater reliability computed for this coding. Risk scores for birth and perinatal complications, such as those found in the Sheffield birth score (Carpenter, Gardner, Jepson et al., 1983) will have to be computed from medical record data.

References

Baghurst PA, McMichael AJ, Wigg NR, et al. (1992). Environmental exposure to lead and children's intelligence at the age of seven years: The Port Pirie cohort study. *New England Journal of Medicine*, 327, 1279–1284.

Behavioral Risk Factor Surveillance Survey (1994). *Morbidity and Mortality Weekly Report*, 43, 375–380.

Blehar MC, Lieberman AF, and Ainsworth MD (1977). Early face-to-face interaction and its relation to later infant-mother attachment. *Child Development*, 48, 182–194.

Bradley RH and Caldwell BM (1980). Home environment, cognitive competence and IQ among males and females. *Child Development*, 51, 1140–1148.

Buka SL and Earls F (1993). Early determinants of delinquency and violence. *Health Affairs*, Winter, 46–64.

Buka SL, Tsuang MT, and Lipsett LP (1993). Pregnancy/delivery complications and psychiatric diagnosis: A prospective study. *Archives of General Psychiatry*, 50, 151–156.

Busch KG, Zagar R, Hughes JH, Arbit J, and Bussell RE (1990). Adolescents who kill. *Journal of Clinical Psychology*, 46, 472-485.

Carpenter RG, Gardner A, Jepson M, et al. (1983). Prevention of unexpected infant death: An evaluation of seven years of birth scoring and increased visiting of high risk infants to prevent unexpected deaths in Sheffield. *Lancet*, i, 723-727.

Caspi A, Henry B, McGee RO, and Moffitt TE (1995). Temperamental origins of child and adolescent behavior problems: From age three to fifteen. *Child Development*, 66, 55-68.

Denno DW (1990). *Biology and Violence: From Birth to Adulthood*, Cambridge, U.K.: Cambridge University Press.

Downdy L, Mrazek D, Quinton D, and Rutter M (1984). Observation of parent-child interaction with 2 to 3 year olds. *Journal of Child Psychiatry and Psychology*, 25, 379-407.

Earls F and Barnes McGuire J (in press). Understanding and preventing child abuse in urban settings. In: McCord J (ed.), *Aggression and Urban Disadvantage*. Harry Frank Guggenheim Foundation, New York.

Earls F, Cairns R, and Mercy J (1993). The control of violence and the promotion of nonviolence in adolescents. In: Millstein S, Peterson A, and Nightingale E (eds.), *Promoting the Health of Adolescents*. New York: Oxford University Press, pp. 285-304.

Earls F and Carlson M (1995). Promoting human capability as an alternative to early crime prevention. In: Farrington D, Wikstrom PO, and McCord J (eds.), *Integrating Motivational and Opportunity-Reducing Crime Prevention Strategies*. National Council of Crime Prevention, Sweden.

Earls F and Jung K (1987). Temperament and home environment characteristics as causal factors in the early development of childhood psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry*, 26, 491-498.

Glueck S and Glueck ET (1934). *One Thousand Juvenile Delinquents*. Cambridge, MA: Harvard University Press.

Gray J (1975). *Elements of a Two-Process Theory of Learning*. New York: Academic Press.

Gray J (1982). *The Neuropsychology of Anxiety: An Enquiry Into the Functions of the Septal-Hippocampal System*. Oxford: Oxford University Press.

Gray J (1987). Perspectives on anxiety and impulsivity. *Journal of Research in Personality*, 21, 493-509.

Gunnar MR, Porter FL, Wolf CM, Rigatuso J, and Larson MC (1995). Neonatal stress reactivity: Predictions to later emotional temperament. *Child Development*, 66, 1-13.

Hack M, Breslau N, Weissman B, and Aram DM (1991). Effect of very low birth weight and subnormal head size on cognitive abilities at school age. *New England Journal of Medicine*, 325, 231-237.

Institute of Medicine (1990). Subcommittee on Nutritional Status and Weight Gain During Pregnancy. *Nutrition During Pregnancy*, National Academy of Sciences, Washington D.C.: National Academy Press.

Jacobson SW, Jacobson JL, Sokol RJ, Martier SS, and Ager JW (1993). Prenatal alcohol exposure and infant information processing ability. *Child Development*, 64, 1706-1721.

Kagan J (1989). Commentary: Does infant boredom reflect intelligence? *Human Development*, 32, 172-176.

Kagan J (1991). Etiologies of adolescents at risk. *Journal of Adolescent Health*, 12, 591-596.

Kagan J and Snidman N (1991). Infant predictors of inhibited and uninhibited profiles. *Psychological Science*, 2, 40-44.

Klein NK, Hack M, and Breslau N (1989). Children who were very low birth weight: Development and academic achievement at nine years of age. *Journal of Developmental and Behavioral Pediatrics*, 10, 32-37.

Kleinman JC and Madans JH (1985). The effects of maternal smoking, physical stature and educational attainment on the incidence of low birth weight. *American Journal of Epidemiology*, 121, 843-855.

Kochanska G (1993). Towards a synthesis of parental socialization and child temperament in early development of conscience. *Child Development*, 64, 325-347.

Laucht M, Esser G, and Schmidt MH (1994). Contrasting infant predictors of later intelligence. *Journal of Child Psychiatry and Psychology*, 35, 649-662.

Loeber R and Dishion TJ (1983). Early predictors of male adult delinquency: A review. *Psychological Bulletin*, 94, 68-99.

Loeber R and LeBlanc M (1990). Toward a developmental criminology. In: Tonry M and Morris N (eds.), *Crime and Justice: A Review of Research* (vol. 12). Chicago: University of Chicago Press, pp. 375-473.

McCall RB and Carriger MS (1993). A meta-analysis of infant habituation and recognition memory performance as predictors of later IQ. *Child Development*, 64, 57-79.

Meaney MJ, Aitken DH, van Berkel C, Bhatnagar S, and Sapolsky RM (1988). Effect of neonatal handling on age-related impairments associated with the hippocampus. *Science*, 239, 766-768.

Moffitt TE (1990). The neuropsychology of juvenile delinquency: A critical review of research and theory. In: Tonry M and Morris N (eds.), *Crime and Justice: A Review of Research* (vol. 12). Chicago: University of Chicago Press, pp. 99-169.

Moffitt TE (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100, 674-701.

Moffitt TE, Gabrielli WF, and Mednick SA (1988). Socioeconomic status, IQ and delinquency. *Journal of Abnormal Psychology*, 90, 152-156.

Myerberg DZ, Carpenter RG, Myerberg CF, Britton CM, Bailey CW, and Fink BE (1995). Reducing postneonatal mortality in West Virginia: A statewide intervention program targeting risk identified at and after birth. *American Journal of Public Health*, 85, 631-637.

National Institute on Alcohol Abuse and Alcoholism (1990). *Alcohol and Health*, Seventh special report to the U.S. Congress from the Secretary of Health and Human Services, Rockville, MD.

Pascoe JM, Chessare J, Baugh E, et al. (1987). Help with prenatal household tasks and newborn birthweight: Is there an association? *Journal of Developmental and Behavioral Pediatrics*, 8, 207-212.

Pascoe JM and Earp JA (1984). The effect of mothers' social support and life changes on the stimulation of the child in the home. *American Journal of Public Health*, 74, 358-360.

Quay HC (1993). The psychobiology of undersocialized aggressive conduct disorder: A theoretical perspective. *Development and Psychopathology*, 5, 165-180.

Raine A, Brennan P, and Mednick SA (1994). Birth complications combined with early maternal rejection at age 1 predispose to violent crime at age 18 years. *Archives of General Psychiatry*, 51, 984-988.

Rasch G (1980). *Probabilistic Models for Some Intelligence and Attainment Tests*, Chicago: University of Chicago Press.

Raudenbush SW and Bryk AJ (1992). *Hierarchical Linear Modeling: Applications and Data Analysis*, Newbury Park, CA: Sage Publications, Inc.

Reiss AJ and Rhodes AL (1961). The distribution of juvenile delinquency in the social class structure. *American Sociological Review*, 26, 720–732.

Rende R (1993). Longitudinal relations between temperament traits and behavioral syndromes in middle childhood. *Journal of the American Academy of Child and Adolescent Psychiatry*, 32, 287–290.

Robins L (1978). Sturdy childhood predictors of adult antisocial behavior: Replications from longitudinal studies. *Psychological Medicine*, 8, 611–622.

Rose SA (1983). Differential rates of visual information processing in full term and preterm infants. *Child Development*, 54, 1189–1198.

Rose SA (1994). Relation between physical growth and information processing in infants born in India. *Child Development*, 65, 889–902.

Rose SA and Feldman JF (1995). Prediction of IQ and specific cognitive abilities at 11 years from infancy measures. *Developmental Psychology*, 31, 685–696.

Rose SA, Feldman JF, and Wallace IF (1992). Infant information processing in relation to six-year cognitive outcomes. *Child Development*, 63, 1126–1141.

Rothbart MK, Ahadi SA, and Hershey KL (1994). Temperament and social behavior in childhood. *Merrill-Palmer Quarterly*, 40, 21–39.

Serdula M, Williamson DF, Kendrick JS, et al. (1991). Trends in alcohol consumption by pregnant women 1985–1988, *Journal of the American Medical Association*, 265, 876–879.

Strassberg Z, Dodge KA, Petit GS, and Bates JE (1994). Spanking in the home and children's subsequent aggression toward kindergarten peers. *Development and Psychopathology*, 6, 445–461.

Straus MA (1979). Measuring intrafamily conflict: The conflict tactics scale. *Journal of Marriage and the Family*, 41, 75–88.

Tronick EZ and Beeghly M (1991). Effects of prenatal exposure to cocaine on newborn behavior and development: A critical review. *OSAP Prevention Monographs*, 11, 25–48.

Ventura SJ, Martin AM, Taffel SM, Matthews TJ, and Clarke SC (1994). Advance report of final natality statistics. *Monthly Vital Statistics Report* 43, No. 5, Supplement, October 25, 1994, Centers for Disease Control and Prevention/National Center for Health Statistics.

Welner Z and Rice J (1988). School-age children of depressed parents: A blind and controlled study. *Journal of Affective Disorders*, 15, 291–302.

White JL, Moffitt TE, Earls F, Robins L, and Silva PA (1990). How early can we tell?: Predictors of childhood conduct disorder and adolescent delinquency. *Criminology*, 28, 507–533.

Wright BD and Masters GN (1982). *Rating Scale Analysis: Rasch Measurement*, Chicago: MESA Press.

Yogman MW, Kindlon D, and Earls F (1995). Father involvement and cognitive/behavioral outcomes of premature infants. *Journal of the American Academy of Child and Adolescent Psychiatry*, 34, 1–13.

Chapter 9. Exposure to Violence

by Stephen L. Buka, Sc.D., and Mary Beth Selner-O'Hagan, Ph.D.

A comprehensive assessment of exposure to violence (ETV) was needed to further the Project's goal of obtaining a better understanding of the causes and pathways of delinquency, crime, substance use, and violence. It has been suggested that ETV during childhood and adolescence is widespread and has an adverse impact on the adaptational functioning of youths, in terms of both the perpetuation of violence and the development of emotional problems. Recent reports indicate that exposure to violent events places youths at risk for engaging in violent acts. For example, investigation of intrafamilial ETV has established an association between abusive parenting and exhibiting violent behavior (Dodge, Bates, and Pettit, 1990; Farrington, 1991; Patterson, 1986; Patterson, Debaryshe, and Ramsey, 1989; Straus, 1991). Further, DuRant and his colleagues (DuRant, Cadenhead, Pendergrast, Slavens, and Linder, 1994) found that ETV in the community was the strongest concurrent predictor of use of violence by adolescents. The adverse impact of ETV is also evident in the psychiatric sequelae of such exposure, including depression, withdrawal, fear, anxiety, affect dysregulation, aggression, dissociative reactions, and intrusive thoughts (Bell and Jenkins, 1990; DuRant, Getts, Cadenhead, Emans, and Woods, 1995; Fantuzzo, DePaola, Lambert, Martino, Anderson, and Sutton, 1991; Fitzpatrick and Boldizar, 1993; Freeman, Mokros, and Poznanski, 1993; Hyman, Zelikoff, and Clarke, 1988; Lyons, 1987; Martinez, Richters, and Benoit, 1993; Osofsky, Wewers, Hann, and Fick, 1993).

One of the more frequently discussed effects of ETV is the development of posttraumatic stress disorder (PTSD; Fitzpatrick and Boldizar, 1993; Pynoos, Frederick, Nader et al., 1987; Pynoos and Eth, 1984). Empirical studies have linked both chronic and acute ETV as a witness or a victim to the symptomatology of PTSD, particularly among younger persons (Dyson, 1990; Pynoos, 1985). Although PTSD is recognized as a consequence of catastrophic exposure to armed conflict or natural disasters, the definition of this disorder has undergone continual refinement and now includes exposure to the types of street and domestic violent events that occur all too routinely in American cities.

Shortcomings of Previous ETV Research

Many authors urge the development of research to understand both the short- and long-term consequences of ETV. Richters (1993) stresses the need to delineate which types of violent experiences in different contexts carry the greatest threat in order to develop and implement appropriate intervention strategies with those at greatest risk for maladjustment. Pynoos et al. (1987) support this view and assert that a child's reaction to trauma is affected by proximity to the violent event, the child's relationship with the victim, and the presence of a parent or caretaker. Putnam and Trickett (1993) note that the heterogeneous nature and severity of traumatic exposure confound this research. Research needs to clarify carefully the nature and types of traumatic stressors so that definitions and criteria can be standardized across studies and instruments can be developed better to quantify levels of exposure. To expand the initial

findings surrounding ETV and to obtain a more accurate understanding of this in terms of type and prevalence of exposure, risk and protective factors that interact with exposure, and the developmental impact of exposure, it will be necessary to use longitudinal designs, draw more representative samples, and improve on existing assessment procedures.

The design of the Project provides a unique opportunity to address several shortcomings of previous ETV research. This project provides a heterogeneous sample of white, African-American, and Hispanic male and female subjects from diverse economic groups who will be followed longitudinally. Although this project affords the setting necessary to comprehensively assess ETV in terms of sampling plan and longitudinal design, existing instruments of ETV were not considered appropriate for detailing exposure. Thus, the Project has worked toward developing an instrument that addresses the inadequacies in existing measures of ETV. In doing so, improvement was sought on existing measures of ETV in terms of:

- Defining what events are considered "violent."
- Specifying whether events are experienced or witnessed.
- Obtaining information on contextual features of the exposure.
- Separating acute from chronic exposure.
- Yielding an interval-level measure that conforms to the requirements of the longitudinal design.
- Providing information on the psychometric properties of the instrument.

Defining violent events. Decisions regarding the content of ETV instruments have varied widely across investigations. Most instruments include questions regarding more severe events, such as shootings, being attacked with a weapon, or seeing someone killed. Some include questions regarding drug-related violence. Few contain questions regarding sexual violence. Some studies include violence that is observed in the media; others include violent events that subjects are told happened to someone else. Such divergence in definitions limits the comparisons that can be made across studies and complicates one's ability to estimate the prevalence of exposure to violent events in the lives of young persons.

Victimization versus witnessing. Although differentiating between violence that is witnessed and violence that is experienced (victimization) is considered to be important in charting distinct sequelae, many studies do not separate such events (Fantuzzo et al., 1991). In addition, despite findings that abused children react differently to witnessed violence than nonabused children, this distinction has rarely been made (Hughes, 1988; Hughes, Parkinson, and Vargo, 1989).

Context of the exposure. Although many studies have inquired about ETV in diverse settings (e.g., in the community versus in the home), most instruments lack the specificity required to make separate assessments as to the nature or impact of exposure in various settings. Characteristics of violent events (e.g., setting, severity, whether witnessed alone or with others) have not been carefully studied, even though they are thought to have a mediating influence on the impact of ETV. For example, separating the effects of community violence from that experienced within the home has not been achieved. Differences in the impact of exposure based on whether the "offender" was someone the subject knew have not been investigated. Further, although identification with victims is thought to be important (e.g., Pynoos et al., 1987), studies have not routinely included questions to discern the relationship between the subject and victim(s).

Acute versus chronic exposure. There has been no systematic comparison of the effects of a traumatic isolated incident with exposure to recurring, chronic violence (i.e., single, sudden, and unexpected versus longstanding and anticipated). Although some investigators have looked at the impact of specific violent events on children, neither the differential impact of experiencing either acute or chronic violence nor the interaction of experiencing both types of events has been thoroughly investigated.

Interval scale. Scores on ETV instruments are usually treated as falling on an interval scale with all items being weighted equally. For example, the most widely used measure of ETV (Richters and Saltzman, 1990) uses a frequency scale that ranges from 0 (no exposure) to 8 (exposed every day). This technique assumes an interval scale that necessarily requires that the increase in ETV impact between 0 (no exposure) and 1 (exposed once) is the same as the increase in ETV impact between 7 (exposed once a week) and 8 (exposed every day). There is no justification for the existence of equal intervals between scale points in the measurement of ETV. Further, most measures weigh violent events equally despite extensive qualitative differences between item content (e.g., seeing someone hit versus seeing someone shot at versus being shot). It seems clear that two persons with very different types of exposure who receive the same score (e.g., someone who has witnessed four murders obtains the same score as someone who has seen someone chased four times) probably are affected differently by these events and that scoring systems should reflect the severity of the violence experienced.

Psychometric properties. Little research has been conducted to assess the psychometric properties of existing instruments of ETV (Gladstein, Rusonis, and Heald, 1992; Richters and Martinez, 1993; Cooley-Quille, Turner, and Beidel, 1995). Cooley-Quille et al. (1995) have begun to assess the reliability and validity of one scale, the Children's Report of Violence (CREV), and have obtained modest results that support use of this instrument in assessment of community violence. The investigation of the psychometric properties of a measure that includes questions regarding family and sexual violence, assesses ETV in separate settings (home, school, and community), and includes other contextual aspects of the exposure has not been initiated.

A New ETV Instrument: My Exposure to Violence

In response to these limitations, the Project has completed a three-phase development sequence to create a new measure of ETV that addresses many shortcomings of prior ETV measures. In the first phase, the Project examined the fit of a set of ETV response data from 49 (29 females, 20 males) low- to middle-income Boston-area residents (age range: 9 to 24 years) to the Rasch model (see chapter 12, "Item Response Theory" [IRT], for a description of this model). The results of this analysis (Kindlon, Wright, Raudenbush, and Earls; in press) gave quite strong support for the conception of ETV as a latent construct. Results indicated that the severity of items, as indicated by higher scores, conformed to researchers' theoretical expectations: the empirically more extreme/severe items correspond to the more severe events (e.g., violence with deadly weapons). Moreover, the scale separated respondents well into a unimodal and symmetric distribution, although there were gaps between items that would optimally be filled by items of intermediate severity. The statistics measuring the fit of the data (for both persons and items) to the model were also quite good. Fit statistics are calculated by comparing predicted item scores with observed item scores. For example, the model predicts that extreme items should not be endorsed by persons with low exposure to violence, but less extreme items should be endorsed by persons with a high level of exposure. The primary fit statistic for a person measure or item calibration is the mean of the standardized square residuals (observed–expected) for persons over an item or items over a person. The expected values for the mean standardized fit statistic is 0 (SD = 1). In these data, the value for item fit was -.2 (SD = .6) and for person fit, -.3 (SD = .6), indicating good fit to the Rasch model. The key weakness of the measure was evident in relatively low reliability (Cronbach alpha = .56), indicating that the measure had too few items of moderate correlation.

The second phase in measure development was to replicate this analysis on a larger (N = 97; 53 females, 44 males; age range 12 to 18 years) and more diverse sample in Chicago to assess whether the initial promising results would generalize to the setting of interest in the proposed study. Results indicated that the data were quite similar in structure to that found in the original study. The hierarchy of items was again consistent with theoretical expectations. Reliability was of a similar magnitude (Cronbach alpha = .70), and item and person fit statistics were quite good. For items, the standardized fit statistic M = -.1, SD = 1.2 and for persons M = -3, SD = .7. In this analysis, unlike the Boston study, there was a small subset of individuals (around 3 percent) whose responses did not fit the model. These were persons who showed "acute" rather than "chronic" exposure, that is, a severe event(s) experienced in the relative absence of less severe events. These individuals constitute a separate group for analysis. Researchers also conducted initial validity studies with these data. As expected, the ETV measure derived from the IRT modeling distinguishes among neighborhoods with respect to homicide rate; higher levels of neighborhood violence are associated with greater individual ETV. Additional evidence for measure validity is provided by the fact that greater ETV was seen for persons who reported having committed a violent offense. African Americans reported higher levels of ETV than either whites or Hispanics. There were no age or gender differences.

In the third phase of development, a field test of a significantly enhanced ETV measure was conducted in Chicago. This instrument, My Exposure to Violence (My ETV), focuses on the subject's exposure to a variety of violent and/or traumatic events over the subject's lifetime and during the past year. It includes questions regarding 24 different violent events that were witnessed or personally experienced by the subject, or that happened to someone the subject knows. An open-ended question regarding exposure to any violent event not covered within the instrument also is included. Frequency of exposure, both during the past year and lifetime, is measured on a 6-point scale (0–5), ranging from "never" to "more than 50 times." Followup questions are included for events that occurred in the past year, specifying the location (home, school, in neighborhood, outside neighborhood) and identity of victim and/or perpetrator (e.g., family member, someone known, stranger). (See sample questions 5 and 6 in questionnaire 9.1.) In addition, a section at the end of the instrument prompts the respondent to select the events that were most upsetting and obtains detailed questions regarding the context and impact of these events (e.g., timing, location, proximity, identity of victims and perpetrators, alone versus with others, fear of death). The primary objectives of this third phase include (1) developing an instrument that uses a scoring system based on an interval scale that will weight a broad range of violent events according to their severity and that differentiates between violence that is directly experienced (victimization) and that is witnessed, and (2) assessing the psychometric properties of this instrument.

Eighty subjects between the ages of 9 and 24 participated in this third phase. All Project study participants who were interviewed from mid-September through early November 1995 were eligible, with the exception of those respondents who had a limited amount of time for the interviews and those who did not speak English. Given that subjects were taken from a consecutive string of enrollees, there was no stratification by neighborhood, gender, or other criteria used in the Project sampling design. Sixty-one percent of the sample was male, and the racial/ethnic composition was varied, with 46 percent African American, 40 percent white, 10 percent Hispanic, and 4 percent other, based on subject self-descriptions. Subjects were distributed across the study cohort ages: (ages 9 and 12: 26 percent; age 15: 19 percent; age 18: 26 percent; ages 21 and 24: 29 percent). Assessment of the neighborhood clusters for this sample indicated that 50 percent resided in low-crime areas, 17 percent in moderate-crime areas, and 33 percent in high-crime neighborhoods.

Thirty subjects were recontacted between 2 and 4 weeks following the initial interview to assess the test-retest reliability of My ETV. Of those contacted, 5 were available only after the 4-week interval and 1 declined to participate due to scheduling conflicts, resulting in 24 subjects participating in the retest.

Three primary scales were derived from the My ETV measure: (1) witnessing violent events, (2) victimization, and (3) total exposure (witnessing plus victimization). Scores were obtained for both lifetime and past-year exposure for each of these, resulting in a total of six scales. Initially, these scales were examined for their fit to a Rasch model. Results indicated that the fit statistics for these six scales were quite good (fit statistic $< |2.0|$). Another way to examine goodness of fit is at the individual level. For persons, poor fit (again defined as a fit

statistic > |2.0|) was rare and in virtually all cases was manifest by an "acute" pattern of ETV. An acute pattern was defined by a fit statistic > |2.0| and a measure of ETV that fell below the mean for the group. In order to exhibit this combination, one would have to endorse one or more of the more severe types of exposure. An example of this pattern was a subject who reported that her only experiences with violence were being the victim of sexual assault on more than one occasion. The percentage of acute subjects (depending on the scale used) ranged between 0 and 5 percent. Lastly, an inspection of the hierarchy of item extremity confirmed the expectation that higher item calibrations would be associated with more extreme events. Two extreme items that did not completely fit these expectations were "home during break-in" and "witness to a sexual assault."

In terms of psychometric properties, both reliability and validity of these six scales were investigated. Cronbach's alpha was used to assess the internal consistency of each of the six scales. Results reflected excellent reliability (.68–.93), with the highest levels obtained on the witnessing (.91–.92) and total (.89–.93) ETV scales. The test-retest reliability for all six scales was also good (.61–.92), with reports of lifetime exposure being somewhat more reliable (.85–.92) than reports of exposure during the past year (.61–.89).

To investigate the validity of My ETV, the data were analyzed in two ways. First, group differences in the measures derived from the Rasch IRT modeling were examined. Results indicated significant differences ($p < .05$) as a function of gender, race, history of violent offending, and neighborhood violence. Second, data were examined at the level of the individual item. One preselected item (had seen someone shot) from the past year and lifetime scales is presented here to provide a more descriptive sense of group differences.

- **Gender.** Male subjects were over 1.5 times more likely than female subjects to have seen at least one shooting during their lifetime (35 percent and 20 percent, respectively) and 4 times more likely to have witnessed a shooting on more than one occasion (28 percent and 7 percent, respectively).

- **Race/ethnicity.** African-American subjects were three times more likely than whites to report having seen someone shot during their lifetime (49 percent compared to 16 percent) and five times more likely (16 percent to 3 percent) to report seeing someone shot within the past year.

- **Violent offending.** Violent offenders were three times more likely to have seen a shooting during the past year (22 percent to 7 percent).

- **Neighborhood level of violent crime.** Subjects in high-crime neighborhoods were 17 times more likely to have witnessed a shooting during the past year than those who resided in a low-crime neighborhood (35 percent to 2 percent), with exposure to this event falling in between (21 percent) in moderate-crime communities.

This third phase of scale development investigated My ETV as a measure of youth exposure to violence developed through Rasch modeling. Initial results indicate that My ETV is a promising measure of exposure to violence. There are a number of criteria that can be invoked to assess the value of an instrument, including the evaluation of reliability, validity, and practical utility. The scales derived from My ETV are psychometrically sound, and the response data fit well to the Rasch model for rating scales. Results of test-retest and internal consistency analyses indicate that this is an acceptably reliable instrument. Assessment of the criterion validity of My ETV also produced promising results. Particularly good evidence for validity of this measure was obtained in the associations demonstrated between scores on My ETV and gender, violent offending, and neighborhood crime. As expected, males reported more exposure than females, violent offenders reported more exposure than nonoffenders, and those living in high-crime areas reported more exposure than those residing in low-crime areas.

Further Investigation

The three phases of investigation have furthered the development of a psychometrically sound and useful instrument for the assessment of ETV. Future research is needed to assess the contribution of contextual factors, to obtain a more detailed understanding of acute versus chronic ETV, and to investigate the contribution of ETV to the development of psychiatric disorders in general and to PTSD in particular. The inclusion of contextual information, such as the location of the exposure and the identity of those involved (perpetrator and victim), and the differentiation of chronic and acute exposure will assist in delineating those events that pose the greatest threat in terms of short- and long-term sequelae. Identifying youngsters prior to exposure and then following their development during and after exposure will assist in more clearly assessing the contribution of these factors to both the determinants of ETV and the resulting impacts. The Project plans to incorporate this new measure in Wave 2, which will lead to the completion of these goals.

References

Bell C and Jenkins E (1990). Preventing black homicide. In: J. Dewart (ed.), *The State of Black America*. New York: National Urban League, pp. 143–155.

Cooley-Quille M, Turner S, and Beidel D (1995). Emotional impact of children's exposure to community violence: A preliminary study. *Journal of the Academy of Child and Adolescent Psychiatry*, 23(10), 1362–1368.

Dodge K, Bates J, and Pettit G (1990). Mechanisms in the cycle of violence. *Science*, 250, 1678–1683.

DuRant R, Cadenhead C, Pendergrast R, Slavens G, and Linder C (1994). Factors associated with the use of violence among urban black adolescents. *American Journal of Public Health*, 84(4).

DuRant R, Getts A, Cadenhead C, Emans SJ, and Woods E (1995). Exposure to violence and victimization and depression, hopelessness, and purpose in life among adolescents living in and around public housing. *Developmental and Behavioral Pediatrics*, 16(4), 233–237.

Dyson J (1990). The effect of family violence on children's academic performance and behavior. *Journal of the National Medical Association*, 82(1), 17–22.

Fantuzzo J, DePaola L, Lambert L, Martino T, Anderson G, and Sutton S (1991). Effects of interparental violence on the psychological adjustment and competencies of young children. *Journal of Consulting and Clinical Psychology*, 59, 258–265.

Farrington D (1991). Childhood aggression and adult violence: Early precursors and later life outcomes. In: DJ Pepler and KH Rubin (eds.), *The Development and Treatment of Childhood Aggression*. Hillsdale, NJ: Erlbaum.

Fitzpatrick K and Boldizar J (1993). The prevalence and consequences of exposure to violence among African-American youth. *Journal of the American Academy of Child and Adolescent Psychiatry*, 32(2), 424–430.

Freeman L, Mokros A, and Poznanski E (1993). Violent events reported by normal urban school-aged children: Characteristics and depression correlates. *Journal of the American Academy of Child and Adolescent Psychiatry*, 32(2), 419–423.

Gladstein J, Rusonis E, and Heald F (1992). A comparison of inner-city and upper-middle class youths' exposure to violence. *Journal of Adolescent Health*, 13(4), 275–280.

Hughes H (1988). Psychological and behavioral correlates of family violence in child witnesses and victims. *American Journal of Orthopsychiatry*, 58, 77–90.

Hughes H, Parkinson D, and Vargo M (1989). Witnessing spouse abuse and experiencing physical abuse: A "double whammy?" *Journal of Family Violence*, 4, 197–209.

Hyman A, Zelikoff W, and Clarke J (1988). Psychological and physical abuse in the schools: A paradigm for understanding post-traumatic stress disorder in children and youth. *Journal of Traumatic Stress*, 1, 243–266.

Kindlon D, Wright B, Raudenbush S, and Earls F (in press). The measurement of children's exposure to violence: A Rasch analysis. *International Journal of Methods in Psychiatric Research.*

Lyons J (1987). Post-traumatic stress disorder in children and adolescents: A review of the literature. In: Chess S and Thomas A (eds.), *Annual Progress in Child Psychiatry and Development*. New York, NY: Brunner Mazel, pp. 451-457.

Martinez P, Richters J, and Benoit M (1993). The NIMH Community Violence Project II: Children's distress symptoms associated with violence exposure. *Psychiatry: Interpersonal and Biological Processes*, 56, 22-35.

Osofsky J, Wewers S, Hann D, and Fick A (1993). Chronic community violence: What is happening to our children? *Psychiatry*, 56, 36-45.

Patterson G (1986). Performance models for antisocial boys. *American Psychologist*, 41, 432-444.

Patterson G, Debaryshe B, and Ramsey E (1989). A developmental perspective on antisocial behavior. *American Psychologist*, 44, 329-335.

Putnam FW and Trickett RK (1993). Child sexual abuse: A model of chronic trauma. *Psychiatry*, 56, 82-95.

Pynoos R (1985). Children traumatized by witnessing of personal violence. In: Eth S and Pynoos R (eds.), *Posttraumatic Stress Disorder in Children*. Washington, D.C.: American Psychiatric Press.

Pynoos R and Eth S (1984). The child as witness to homicide. *Journal of Social Issues*, 40, 87-108.

Pynoos R, Frederick C, Nader K, et al. (1987). Life threat and post traumatic stress in school-age children. *Archives of General Psychiatry*, 44, 1057-1063.

Richters J (1993). Community violence and children's development: Toward a research agenda for the 1990s. *Psychiatry*, 56.

Richters J and Martinez P (1993). The NIMH Community Violence Project I: Children as victims of and witnesses to violence. *Psychiatry: Interpersonal and Biological Processes*, 56, 7-21.

Richters J and Saltzman W (1990). *Survey of Children's Exposure to Community Violence: Parent Report*. Rockville, MD: National Institute of Mental Health.

Straus M (1991). Discipline and deviance: Physical punishment of children and violence and other crime in adulthood. *Social Problems*, 38, 133-154.

Questionnaire 9.1

Sample Items from My Exposure to Violence

5. In your whole life, have you **EVER** seen *someone else* get attacked with a weapon, like a knife or bat? Would you say . . . (read all choices)

1.	Yes	Go to Q. 5A	
2.	No	Go to Q. 6	5 ☐ ____

5A. About how old were you the **first time** you saw someone get attacked with a weapon? (Record age)

5A ☐ ____ (yrs old)

5B. Now, just thinking about the **PAST YEAR**, about how many times have you seen *someone else* get attacked with a weapon, like a knife or bat? Would you say . . . (read all choices)

0.	Never	Go to Q. 6	
1.	Once	Go to Q. (5D)	
2.	2 or 3 times	Go to Q. (5D)	5B ☐ ____
3.	4 to 10 times	Go to Q. (5D)	
4.	More than 10 times	Go to Q. (5C)	

5C. About many times have you seen this **in the PAST YEAR**? (Record number of times→)

5C ☐ ____ (times)

(5D.) When you saw someone get attacked with a weapon **in the PAST YEAR**, did this happen. . .

5D. inside your home?	1. Yes 2. No	5D ☐ ____
5E. in someone else's home?	1. Yes 2. No	5E ☐ ____
5F. at school?	1. Yes 2. No	5F ☐ ____
5G. somewhere else in your neighborhood?	1. Yes 2. No	5G ☐ ____
5H. outside your neighborhood?	1. Yes 2. No	5H ☐ ____

(5I.) **In the PAST YEAR**, did you see someone else attacked with a weapon. . .

5I. by someone in your family?	1. Yes 2. No	5I ☐ ____
5J. by someone else you knew?	1. Yes 2. No	5J ☐ ____
5K. by anyone else?	1. Yes 2. No	5K ☐ ____

(5L). **In the PAST YEAR**, who did you see get attacked with a weapon? Were they. . .

5L. someone you knew and felt very close to?	1. Yes 2. No	5L ☐ ____
5M. someone you knew but didn't feel very close to?	1. Yes 2. No	5M ☐ ____
5N. anyone else?	1. Yes 2. No	5N ☐ ____

5 = DU 6 = REF 7 = NAP 8 = DK 9 = NA

6. In your whole life, have you **EVER** been attacked with a weapon, like a knife or bat? Would you say. . . (read all choices)

0.	Never	Go to Q. 7
1.	Once	Go to Q. 6B
2.	2 or 3 times	Go to Q. 6B
3.	4 to 10 times	Go to Q. 6B
4.	More than 10 times	Go to Q. 6A

6 ☐ ____

6A. About how many times has this happened **in the PAST YEAR**? (Record number of times→)

6A ☐ ____ (times)

6B. About how old were you the **first time** you were attacked with a weapon? (Record age→)

6B ☐ ____ (yrs old)

5 = DU | 6 = REF | 7 = NAP | 8 = DK | 9 = NA

Chapter 10. Heart Rate as a Physiologic Marker

by Daniel Kindlon, Ph.D.

Beginning in 1992, Project staff scientists Enrico Mezzacappa and Daniel Kindlon, under the direction of Felton Earls and in collaboration with Richard Tremblay at the University of Montreal, have been investigating the association between heart rate (HR) in childhood and adolescence and concurrent or subsequent antisocial behavior. These investigations have had two primary facets. The first of these has been to explore emerging technology and methodology for the measurement of heart rate; the second has been to study heart rate as a physiologic marker within a longitudinal study of children in Montreal, Quebec, Canada.

The initial phases of these investigations involved consultation with leading cardiologists and physiological researchers to understand the limitations of simple pulse rate as a physiologic marker. The primary limitation with using pulse rate is that it does not allow for a disentangling of the autonomic nervous system influences that determine the overall rate at which the heart beats and the variability in that rate over time. Spectral analytic techniques, however, where the beat-to-beat variability in heart rate can be separated into sympathetic and parasympathetic influences, do afford a more promising technology. These initial investigations resulted in the publication of a paper (Mezzacappa, Kindlon, Earls, and Saul, 1994) that describes the "how to" of spectral analysis and its usefulness in the study of the physiologic contributors to a propensity for antisocial behavior.

These investigators have continued to explore both assessment techniques and emerging technologies. In particular, a protocol has been used wherein electrocardiogram (ECG) and respiratory time series data (inductance plethysmography) are collected from subjects breathing at a specific frequency (15 breaths per minute) in both supine and standing positions. This protocol not only allows for the decomposition of HR variability into sympathetic and parasympathetic components (low and high power), but also permits the study of other aspects of autonomic influence such as transfer magnitude, transfer phase, and coherence. In addition, advances in spectral analytic software and portable ECG equipment have been monitored and incorporated into ongoing work. Figure 10.1 illustrates how spectral analysis separates the HR time series signal into its component frequencies.

The protocol is designed to elicit HR variability from postural and respiratory sources. It is expected that under the experimental conditions employed, linear relationships would exist between the respiratory and the HR variability. It also is assumed that, in the absence of respiratory pathology, cardiac pathology affecting the sinus node and associated conduction tissue, or neuropathy affecting the central or peripheral autonomic nervous systems, any differences in transfer function estimates would reflect differences in central autonomic processing of respiratory–HR relations.

Physiologic assessment takes place in a laboratory setting familiar to the subject. A surface ECG (lead II) and a respiratory time series (inductance plethysmography capturing both

thoracic and abdominal excursions) are obtained. Gain for respiratory leads is calibrated using a breathing apparatus of known volume.

Subjects are first placed in a supine position and trained to breath in synchrony with a metronome at a frequency of 15 breaths per minute (0.25 hertz) to isolate the HR variability generated by respiratory change from that arising in response to postural change, which naturally occurs at around 0.08 to 0.1 hertz, or about 6 times per minute. After a practice period, data are collected for 4 minutes in the supine position. Subjects are then asked to stand quietly while continuing to breath at the same frequency. Another 4 minutes of data are collected in this position.

Processing of Physiologic Data

The interbeat (R-wave to R-wave) intervals of the surface ECG are digitized to within 1 millisecond. The beat series are visually inspected for artifacts. These are manually corrected using accepted methods designed to maintain the integrity of the time series.

The respiratory signal is digitized at 5.0 hertz. HR time series are reconstructed from the R–R interval series at 5.0 hertz to match the respiratory time series. From these two time series, corresponding 2-minute segments are chosen from each of the supine and standing postures. Power spectra and transfer function estimates are generated from linearly detrended time series.

Two frequency bands are identified in each power spectrum: from 0.04 to 0.15 hertz and 0.20 to 0.30 hertz. These are labeled POST power (postural source), and RSA power (respiratory sinus arrhythmia), respectively, to reflect their presumed origins. Estimates of respiratory transfer functions are taken at 0.25 hertz, or at the peak breathing frequency when this deviates slightly from 0.25 hertz. When coherence of the transfer function is less than 0.5 (range 0.0 to 1.0), estimates of transfer magnitude and transfer phase are deemed statistically unreliable and eliminated.

When known sources of the HR variability are quantified concomitantly with HR, the specific transfer of variability from the source to the HR variability can be estimated. Using spectral techniques, this transfer of variability portrays the frequency dependent distribution of the shared variance between the two processes (cross-spectral density), the magnitude of the HR response to corresponding changes in the sources (transfer magnitude), such as lung volume (bpm/liter) or blood pressure (bpm/mm Hg), the timing of the HR responses to changes in the sources (transfer phase in degrees), as well as estimates of the coherence or linearity and statistical reliability of the relationship between the source and the HR response. Graphic illustrations are presented in figure 10.2.

The first paper (Kindlon, Tremblay, Mezzacappa, Earls, Laurent, and Schaal, 1995) emerging from the Project that explores heart rate as a physiologic marker for aggression reports two primary findings. The first is that individual differences in resting heart rate are

moderately stable over 1- and 2-year intervals in middle to late childhood, thus supporting its usefulness in longitudinal research. This study also finds that fighting behavior was related to HR for 11-year-old boys in a cohort sampled from the general population and for 12-year-old boys in one of the two cohorts of boys with disruptive behavior problems. This relationship holds even when controlling for body size and pubertal status. This set of findings is, to the researchers' knowledge, the first reported demonstration of a relationship between HR and physical aggression within a relatively homogeneous lower socioeconomic status sample.

These cohorts continue to be followed into adolescence using the more sophisticated technologies described above. Results appear in a paper under review (Mezzacappa, Tremblay, Kindlon, Saul, Arseneault, Seguin, Phil, and Earls; in press).

References

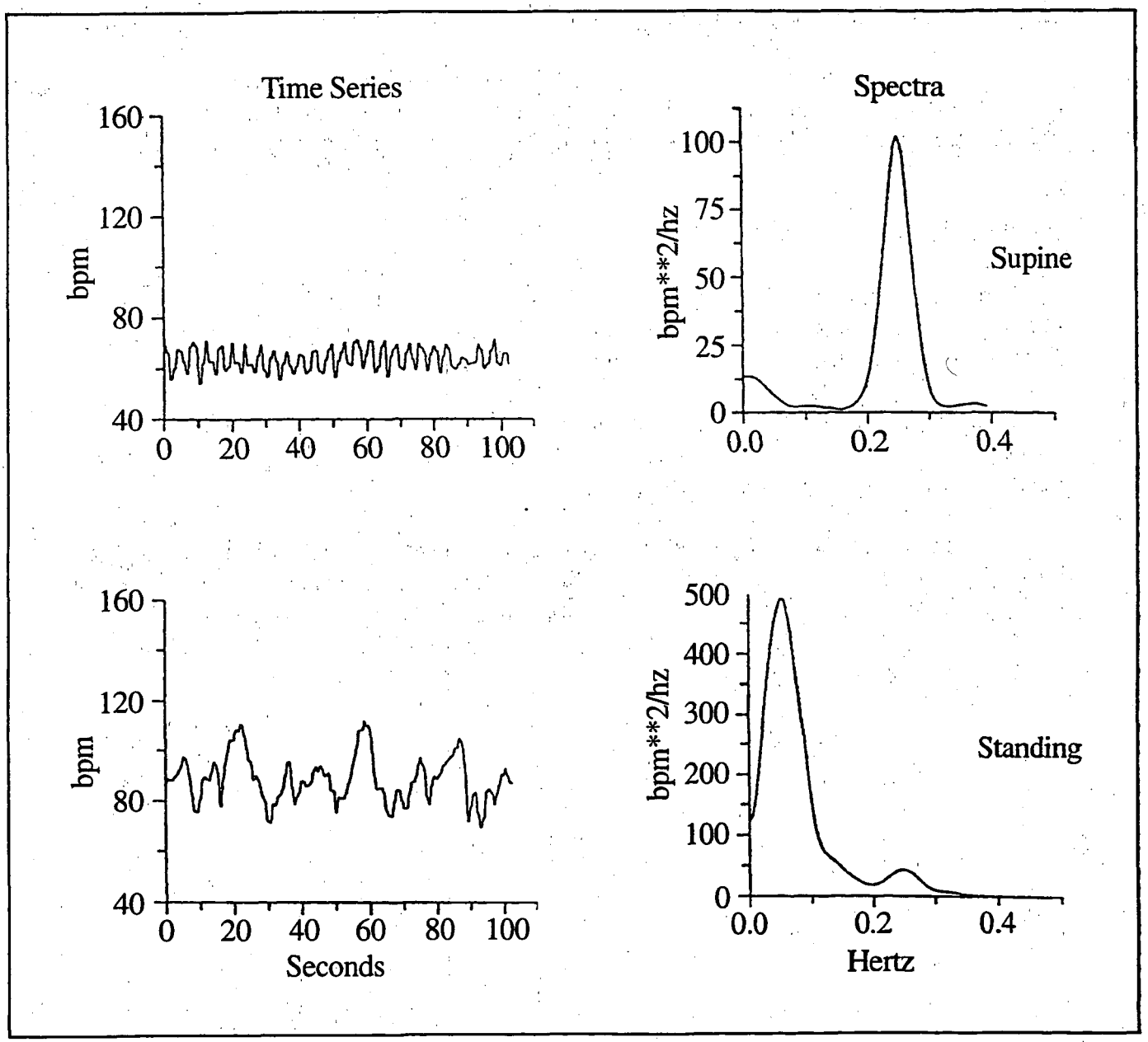
Kindlon DJ, Tremblay RE, Mezzacappa E, Earls F, Laurent D, and Schaal B (1995). Longitudinal patterns of heart rate and fighting behavior in 9- through 12-year-old boys, *Journal of the American Academy of Child and Adolescent Psychiatry*, 34, 371–377.

Mezzacappa E, Kindlon D, Earls F, and Saul JP (1994). The utility of spectral analytic techniques in the study of the autonomic regulation of beat-to-beat heart rate variability. *International Journal of Methods in Psychiatric Research*, 4, 29–44.

Mezzacappa E, Tremblay RE, Kindlon DJ, Saul JP, Arseneault L, Seguin J, Phil RO, and Earls F (in press). Anxiety, antisocial behavior and heart rate regulation in adolescent males. *Journal of Child Psychology and Psychiatry and Allied Disciplines*.

Figure 10.1

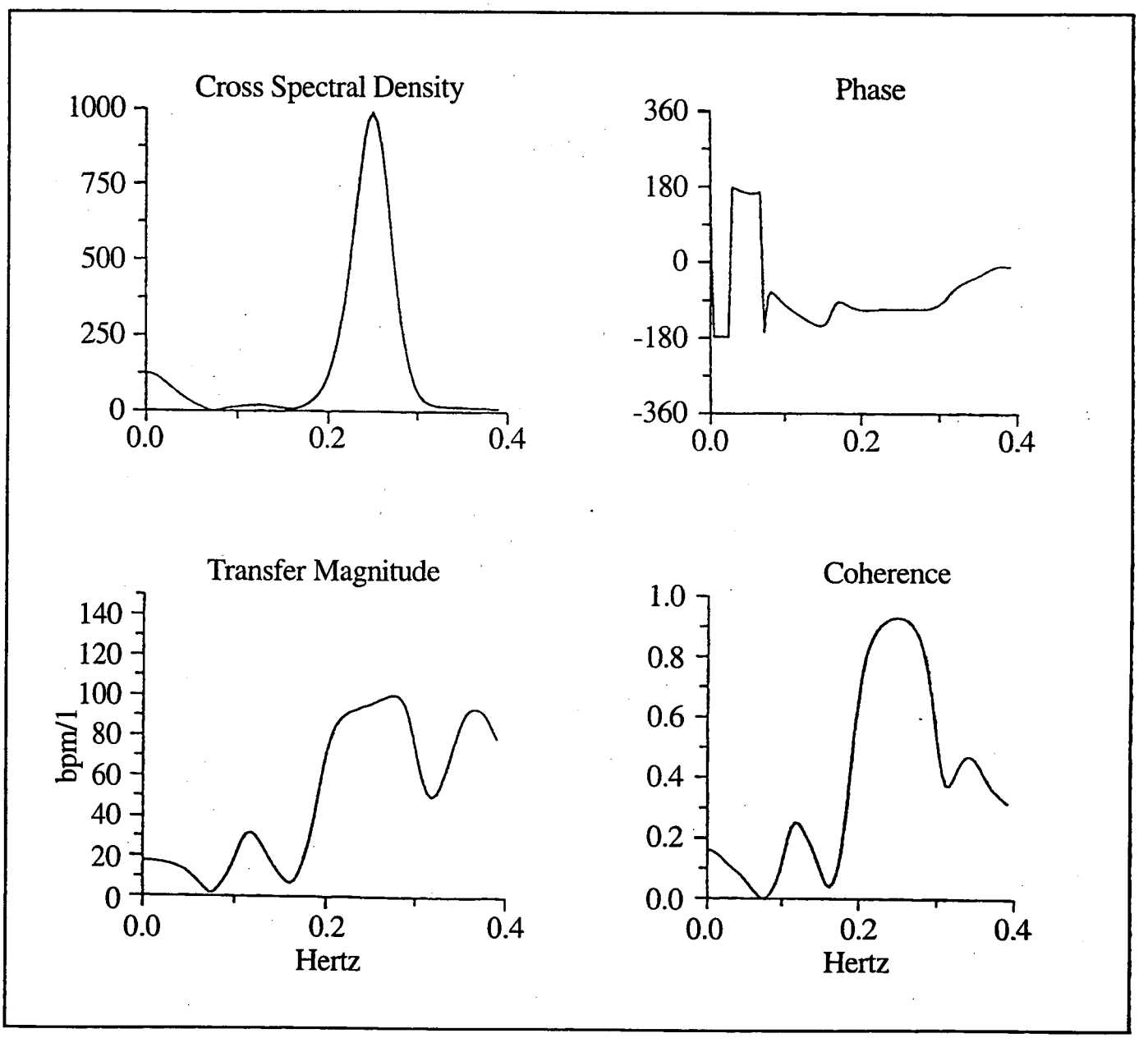
Spectral Analysis of Heart Rate Time Series



Heart rate time series and power spectra in a subject breathing at 0.25 hertz. Note the oscillations of the supine time series and the power spectrum variability around 0.25 hertz. Note the additional slower oscillation in the standing time series and the corresponding variability reflecting blood pressure modulation.

Figure 10.2

Respiration to Heart Rate Transfer Functions



Transfer function estimates from the same subject while in the supine posture.
The region of interest for these indices is that of RSA, from 0.20 to 0.30 hertz.

Part Three: Analytic Strategies

Chapter 11. Hierarchical Linear Models and Growth Models

by Stephen Raudenbush, Ed.D.

A continuing vital task in the planning and conduct of the Project has been the development, testing, and refinement of methods of statistical analysis of the data yielded by the study. Ongoing analytic work has ensured that essential Project hypotheses are testable, sharpened thinking about what the Project can learn from its data, informed key decisions about study design, and produced new generalizable knowledge about the design and analysis of longitudinal studies conducted in the social settings of families, neighborhoods, and schools.

Thinking about statistical modeling is rooted in the conceptual framework of the overall Project. Each child may be viewed as developing along a path that moves away from or toward specific behaviors, and this path is influenced by personal and family characteristics and by neighborhood, school, and peer influences. It is essential that the statistical models employed in this study capture the essential features of this conception, which include notions of "developmental paths," variability of such paths across persons, and the dependence of such paths on personal background and social experience. Such models are referred to as "models for individual change in social context." Raudenbush (1992, 1994, 1995a, 1995b) describes this family of models in some detail and provides numerous examples of their application.

The Project's work on statistical modeling over the past 5 years has included a review of relevant analytic methods, identification of new models and estimation procedures, evaluation of study design alternatives, and initial analysis of the Project's data.

Review of Analytic Approaches

The research design for the Project's cohort data involves a two-stage sampling process: 80 neighborhoods are first sampled; then, within each neighborhood, persons are sampled and followed annually.[1] This design will therefore produce repeated measures on persons who are nested within neighborhoods. Review of current practice revealed that commonly used analytic techniques in the fields of psychology, sociology, criminology, and education are not adequate to meet the demands posed by such longitudinal studies of nested samples. In the rarefied case of continuous outcomes and a perfectly balanced design (equal numbers and spacing of time points for each participant, equal numbers of participants in each

[1]In detail, the sampling plan for the longitudinal cohort study is considerably more complicated, involving the subdivision of Chicago into 343 neighborhoods, the classification of those neighborhoods into 21 strata defined by the cross-classification of ethnic mix and socioeconomic status, the sampling of neighborhoods within strata, the listing of dwelling units in sampled blocks within neighborhoods, the random assignment of listed dwelling units into replicates, and the recruitment of eligible participants within selected replicates. However, for purposes of statistical modeling, there are three key sources of random variation: variation between neighborhoods, variation between participants within neighborhoods, and variation within persons over time.

neighborhood, no missing data, and no covariates), the classical univariate and multivariate analysis of variance will be appropriate to such data. However, the data yielded by this study will include discrete as well as continuous outcomes; it will inevitably produce unequal spacing of time points and some missing data, regardless of the care taken in data collection, and covariates that vary over time, persons, or neighborhoods. Raudenbush (1993a) showed how a family of hierarchical models, estimated via restricted maximum likelihood, can duplicate the results of conventional analyses of variance under the restrictive conditions when those analyses are appropriate, but then extend application to allow for the more complex data that arise in field settings. An important aspect of the Project's analytic work has involved the tailoring of these hierarchical models to deal with a variety of interesting applications that will arise in the analysis of the data. These are briefly considered below.

New Methods and Estimation Procedures

Models for migration across social context. When each person who is repeatedly observed is nested within one and only one social context (e,g., one community), the data can be analyzed by means of a three-level hierarchical linear model. However, if some persons migrate across social contexts (e.g., persons move to a new community, children change classrooms as they progress through school), the data will have a crossed structure. For example, observations may be identified by the cross-classification of persons and communities. Raudenbush (1993) developed maximum likelihood and empirical Bayes estimates for this type of model and applied it to study teacher effects on the language acquisition of children in bilingual education programs during grades one through four. He employed the same model to study the joint effects of neighborhoods and schools on educational attainment in a sample of Scottish school-leavers. Both examples have direct application to the Project's data. Kang (1992), working under Raudenbush's direction, completed a thesis that developed a model including covariance components associated with rows, columns, and cells of unbalanced two-way matrices. This model is applicable to the same types of data structures described above when there is sufficient within-cell data to distinguish the row-by-column interaction variance from the within-cell variance.

Nonnormal error models. Kamali (1992), working under Raudenbush's direction, completed a thesis evaluating the adequacy of four proposed statistical algorithms for nonnormal repeated measures data. Chan (1995) completed a dissertation that developed an algorithm and computing program that will produce maximum likelihood and empirical Bayes estimates for repeated measures on Poisson data. The method uses Monte Carlo integration in combination with rejection sampling to maximize likelihood.

Repeated measures on multivariate outcomes. One stream of Project analyses will involve multivariate repeated measures over time. Thus, a child will have two trajectories of change for a given dependent variable, one for each reporter of that child's social behavior. For example, a parent and a teacher will report on social behavior at each wave of a longitudinal study, yielding two change trajectories that will be related but not identical. This approach is explained in Raudenbush, Brennan, and Barnett (1995); an application to psychological

distress in married couples is in Barnett, Raudenbush, Brennan, Pleck, and Marshall (1995). A related multivariate application in research synthesis ("meta-analysis") is in Kalaian and Raudenbush (in press). This allows synthesis of findings of many studies where not all studies use the same set of multiple outcome variables.

Multilevel structural equation models for latent variables. Not all models for longitudinal data involve repeated measures. Often, the goal will be to relate status on a given developmental indicator to status on other indicators measured later in development. At this time, it is challenging to model these structural relations in the setting of multilevel data. Finding maximum likelihood estimates for such models when persons are nested within neighborhoods and when the numbers of persons vary across neighborhoods has been difficult. Raudenbush (1995c) showed how this problem can be solved in a fairly simple way by conceiving it as a missing data problem and applying the EM (expectation maximization) algorithm.

Collateral work. All of the above work was directly supported by the Project and linked to design and analytic issues in that study. Collateral work that has direct application to the Project involves estimation and interpretation in studies of school effects (Raudenbush and Willms, in press).

Design Alternatives

One of the challenging problems facing the Project has been determination of sample sizes: How many time points should be collected per cohort participant? How many participants should be sampled per neighborhood? How many neighborhoods are needed? How many years should separate the several birth cohorts, and how many cohorts should there be? These questions are not only of keen scientific interest, they are critical to the efficient allocation of resources available to conduct the study. Using the family of hierarchical models described above as a foundation, Project researchers have rigorously examined these questions and, in the process, have generated some new insights available to other researchers using similar designs.

Optimal allocation in longitudinal and nested designs. The researchers have intensively studied the problem of choosing the number of time points per subject, the number of subjects per community, and the number of communities in longitudinal and nested designs. The results are based on variance components analysis from the National Youth Survey. Given these variance components and estimates of the cost of sampling at each level, one can express the sampling variance of the estimator of each fixed effect in the model as a function of the sample sizes at each level. One can then minimize this sampling variance with respect to these sample sizes. The results provide solid support for the design chosen by the Project. Raudenbush (in press) provides the first in a series of papers that report these results.

Linking of data across cohorts. Raudenbush and Chan (1992) pilot tested a new methodology for linking adjacent cohorts in longitudinal accelerated designs. They developed

an empirical test of the credibility of the link and evaluated the power of the test. Raudenbush and Chan (1993) provided evidence on statistical power and multicollinearity associated with alternative designs varying in terms of the number of overlapping time points between adjacent cohorts. That paper also considered the reliability of estimates of individual growth parameters when missing data are present and the consequences of the resulting unreliability for estimation of the fixed effects of the model.

Current Analytic Work

Intensive analysis of the Project's community data has given rise to new methodological work in two areas. First, researchers have developed a multivariate random effects model for count data with two applications. One application involves cross-sectional data on multiple indicators of crime. Researchers are examining variation across neighborhoods in terms of different types of homicide and asking whether community-level structural characteristics have similar or different relationships to the homicide rate, depending on which kind of homicide is of interest. Another application involves the repeated measurement of crime. Researchers are examining community-specific trajectories in homicide rates between 1970 and 1990 and seeking to understand how changes in community structure relate to changing homicide rates. The models are based on a family of Poisson-normal mixture models with the possibility of overdispersion.

The second line of work involves the study of responses to the community survey. Researchers are asking about the psychometric properties of measures of social organization at the neighborhood level, and they are interested in the reliability of variation between persons within neighborhoods. The model being developed handles missing data at the item level. An item response model calibrates each scale, and this model is embedded within a random effects model that estimates variation and covariation of the "true scores" at each of two levels: within and between communities. Based on this work, researchers will be able to generate model-based imputations of missing responses and examine the sensitivity of results to the imputation procedure.

Programming

Richard Congdon has done all of the programming for the new statistical algorithms used in these papers. Additionally, he has continued to develop algorithms and programs for discrete outcomes in multilevel settings having two levels, three levels, and cross-classifications. He continues to work on improving their computational efficiency in anticipation of the large amount of data to be yielded by the community survey and the cohort study.

References

Barnett RC, Raudenbush SW, Brennan RT, Pleck JH, and Marshall NL (1995). Change in marital experiences and change in psychological distress: A longitudinal study of dual-earner couples. *Journal of Personality and Social Psychology*, 69, 5, 839–850.

Chan WS (1992). *Maximum Likelihood Estimation for a Poisson Regression Model with Fixed and Random Effects with Applications in Education.* Proposed dissertation, College of Education, Michigan State University.

Kalaian HA and Raudenbush SW (in press). A multivariate mixed linear model for meta-analysis. To appear in *Psychological Methods*.

Kamali M (1992). *A Study of Single- and Multi-level Logistic Regression Models Using Real and Computer Simulated Data*. College of Education, Michigan State University.

Kang SJ (1992). *A Mixed Linear Model with Two-way Crossed Random Effects and Estimation via the EM Algorithm*. College of Education, Michigan State University.

Raudenbush SW (1992). Hierarchical models for studying school effects and schooling: The next wave of innovation. In: Creemers BPM and Reezigt G (eds.), *Evaluation of Educational Effectiveness*, chapter 5. Enschedde, the Netherlands: Inter-University Consortium on Educational Evaluation and Educational Effectiveness.

Raudenbush SW (1993a). Hierarchical linear models as generalizations of certain common experimental design models. In: Edwards L. (ed.), *Applied Analysis of Variance in Behavioral Science*. New York: Marcell Decker.

Raudenbush SW (1993b). A crossed random effects model for unbalanced data with applications in cross-sectional and longitudinal research. *Journal of Educational Statistics*, 18, 4, 321–349.

Raudenbush SW (1994). Modeling individual and community effects over time: Multilevel statistical models. To appear in Wikstrom PO (ed.), *Integrating Individual and Ecological Aspects on Crime*. Stockholm: National Council on Crime Prevention.

Raudenbush SW (1995a). Statistical models for studying the effects of social context on individual development. In: Gottman J (ed.), *The Analysis of Change*. Hillsdale, NJ: Lawrence Erlbaum.

Raudenbush SW (1995b). Re-examining, reaffirming, and improving application of hierarchical models. *Journal of Educational and Behavioral Statistics*, 20, 2, 210–220.

Raudenbush SW (1995c). Maximum likelihood estimation for unbalanced multilevel covariance structure models via the EM algorithm. *British Journal of Mathematical and Statistical Psychology*, 48, 359-370.

Raudenbush SW (in press). Statistical analysis and optimal design for cluster randomized trials. To appear in *Psychological Methods*.

Raudenbush SW, Brennan RT, and Barnett RC (1995). A multivariate hierarchical model for studying psychological change within married couples. *Journal of Family Psychology*, 9, 2, 161-174.

Raudenbush SW and Chan WS (1992). Growth curve analysis in accelerated longitudinal designs. *Journal of Research on Crime and Delinquency*, 29, 4, 387-411.

Raudenbush SW and Chan WS (1993). Application of a hierarchical linear model to the study of adolescent deviance in an overlapping cohort design. *Journal of Clinical and Consulting Psychology*, 6, 5, 794-806.

Raudenbush SW and Willms JD (in press). The estimation of school effects. To appear in the *Journal of Educational and Behavioral Statistics*.

Chapter 12. Item Response Theory

by Daniel Kindlon, Ph.D., and Stephen Raudenbush, Ed.D.

Item response theory (IRT) has been applied by the Project to create measures that are appropriate for its longitudinal design. Nearly all of the individual-level variables used in social science research are derived from scales with an assumed linear or interval level of measurement; that is, the distance between points on a scale is considered to be equal. A common example is an item with a 3-point response scale defined by 0 (not true), 1 (somewhat true), and 2 (very true). Here the distance between 0 and 1 is assumed to be the same as the distance between 1 and 2. It follows that when items such as these are summed to yield a scale score, the distance between 0 and 10 is assumed to be equal to the distance between 20 and 30. These assumptions are rarely justified, and their violation has unfortunate consequences for longitudinal work such as the Project. When the nonlinear scores (e.g., mean scale score) of two groups, or of persons over time, are compared, quantitative inferences about differences between them cannot be made. The magnitude of the score differences will depend on the characteristics of the items composing the scale. Specifically, the differences will depend on the item probability thresholds (i.e., the level at which the item has a 50-percent probability of being endorsed) and on the spread of these item thresholds (Wright and Masters, 1982).

To overcome some of this, Project researchers investigated the use of IRT (Mellenbergh, 1994; Thissen and Steinberg, 1988). IRT affords a methodology for analyzing the properties of scales, including the modeling of responses as falling on an interval level. Thus, through IRT, interval-level measures can be created from nonlinear raw data if the construct under investigation has the necessary properties.

One initial use of IRT has been with the exposure to violence (ETV) construct. ETV is conceptualized as a condition in which people vary along a continuous distribution ranging from 0 to infinity. Researchers postulate that interval-level measurement is possible along this continuum. They also posit that the types of violent events to which a person could be exposed also vary along the continuum. Some events are more severe than others, and the measurement strategy must capture these differences. The researchers hypothesize that responses to items will be governed by a single latent variable, the degree of exposure to violent events. With this perspective, differences between individuals are not random. For reasons such as area of residence or lifestyle, people who are exposed to one violent event are more likely than others to be exposed to other violent events. It follows that if persons vary on the frequency of their exposure, those with the highest levels of exposure will be most likely to experience the more rare, extreme events. If the ETV construct meets these assumptions, researchers should be able, using IRT, to create an interval-level ETV scale.

The researchers chose to investigate the utility of one of the most straightforward IRT models, the Rasch model for rating scales (Wright and Masters, 1982; Rasch, 1980), with a

set of response data from a short measure of ETV. If these investigations prove fruitful, plans to extend these investigations to other measures in the protocol are anticipated.

The development of the model begins with the question, "What are the important conditions that bear on how an item will be answered?" Rasch modeling is often used in cognitive testing where primary parameters are usually referred to as "person ability" and "item difficulty." For the psychological condition being investigated, these terms are awkward; "item extremity" will replace "item difficulty" to indicate that the more extreme an item is, the less often it is endorsed, and that more extreme items should correspond to more severe forms of violence. "Person exposure" will replace "person ability" to denote the cumulative level of exposure to violence.

With exposure and extremity as the parameters determining the probability of endorsement, a mathematical model can be specified. Since exposure and extremity must be on the same scale to be useful, their natural mathematical relationship is their difference. When item extremity exceeds person exposure, which would be the case when a person with a low exposure is confronted with an item asking whether a severe form of violence has ever been experienced, the probability of endorsement is less than .5. When person exposure is greater than item extremity, the probability is greater than .5. When item extremity and person exposure are equal, the probability of a correct response is .5.

The Rasch model is constructed on a linear scale for ETV by using response data as realizations of the probabilities of item endorsement given the level of that item's extremity and the exposure of the person responding to the item. Since the difference between exposure and extremity can vary from minus infinity to plus infinity, whereas probability varies between 0 and 1, the measure differences are transformed into the 0–1 interval by the natural log-odds ratios. Thus, in the Rasch model, the probability that a person will endorse an item (i.e., answer "yes" where yes = 1 and 0 = no), given item extremity and person exposure, is:

$$\log\left(\frac{P_{ni1}}{P_{ni0}}\right)=B_n-D_i$$

where:
B_n = person exposure
D_i = item extremity
n = persons 1, 2, ..., n
i = items 1, 2, ..., i

In the case of rating scales where there are more than two response options, the equation becomes:

$$\log\left(\frac{P_{nij}}{P_{nij-1}}\right)=B_n-D_i-F_{ij}$$

where:
F_{ij} = step resistance threshold of step j
j = responses 0, 1, 2, ..., j

The step resistance threshold is the adjustment in item extremity that corresponds to making the step between the two points on the response scale being considered for that item (e.g., giving an answer corresponding to a score of 2 rather than of 1). Thus, the probability that is computed refers only to that step.

The units of the log-odds linear scale are called logits, and a person's exposure in logits is defined as his or her natural log-odds for endorsing items of a standard level of extremity. The level of an item's extremity in logits is its natural log-odds for eliciting endorsement from persons with a standard exposure.

Both the preliminary results with pilot data (Kindlon, Wright, Raudenbush, and Earls, in press) and field trials with the Project's ETV measure in Chicago (Selner-O'Hagan, Kindlon, Buka, and Earls, in preparation) have shown that the ETV data fit the Rasch model quite well. These results have encouraged the researchers to use this methodology with other individual-level variables in the Project protocol.

References

Kindlon D, Wright B, Raudenbush SW, and Earls F (in press). The measurement of children's exposure to violence: A Rasch analysis. *International Journal of Methods in Psychiatric Research.*

Mellenbergh GJ (1994). Generalized linear item response theory. *Psychological Bulletin,* 115, 300–307.

Rasch G (1980). *Probabilistic Models for Some Intelligence and Attainment Tests.* Chicago: University of Chicago Press.

Selner-O'Hagan MB, Kindlon D, Buka SL, and Earls F (submitted). Assessing exposure to violence in urban youth.

Thissen D and Steinberg L (1988). Data analysis using item response theory. *Psychological Bulletin,* 104, 388–395.

Wright BD and Masters GN (1982). *Rating Scale Analysis: Rasch Measurement.* Chicago: MESA Press.

Part Four: Data Management and Quality Control

Chapter 13. The Data Management System

by Jan D. Dunham, Ph.D.

The task of managing data for a study of the size and scope of the Project is monumental. The volume of data to be collected over 8 years is enormous, as they come from a variety of sources and from different levels of aggregation. Further, the data must conform to high-quality standards and be held in strict confidence to protect the confidentiality of the study participants. The Data Management and Quality Control (DMQC) unit at the Project has designed and created a data management system (DMS), consisting of several interrelated components, to handle the particular challenges of this research study. This group also designed and installed a local area network (LAN) in the Chicago office, established an Internet connection to facilitate remote communication, and developed quality control and data security procedures. In addition to continuing work on DMS components, DMQC is beginning to work with data from several completed surveys and will implement a plan to link these data and agency record data to data from the cohort assessments. DMQC is working to standardize data set preparation procedures to streamline the process and ensure that all data sets can be linked. DMQC also is documenting these efforts so that other researchers may benefit from their experiences.

Local Area Network

A LAN consisting of three servers was designed by DMQC to: (1) provide maximum security for confidential data, (2) support remote communications, (3) provide sufficient expansion capability, and (4) be easy to use, maintain, and upgrade. The network was designed to minimize processing bottlenecks and to maximize the protection of study data by isolating it from programs, files, and the communications server. Study data and applications are protected by disk duplexing. In addition, all data and programs are backed up to tape every day. Backup tapes are transferred to offsite storage once a week. DMQC also provides technical support to Project staff to resolve user problems ranging from printer jams to major software incompatibilities.

Quality Control/Data Preparation

DMQC has developed and implemented procedures for quality control and data preparation for incoming data. Among the quality checks that DMQC has carried out on existing data are: (1) comparing hard copy records to electronic data files to ensure accuracy of keying/scanning, (2) checking for out-of-range values and internally inconsistent values to ensure instruments are being administered correctly, and (3) reviewing codes for open-ended questions to ensure the accuracy of coding. DMQC monitors the quality of data on an ongoing basis and provides feedback to field staff when appropriate. DMQC also ensures that data are linked at the appropriate level. In some cases, e.g., the key informant data and the community indicators data, this means linking at some level of aggregation such as the

census tract or neighborhood cluster. In other cases, such as systematic social observation data and agency records data, it means linking at the level of individual cohort members.

The Data Collection System

The development of a computerized data collection system, which enables research assistants to conduct interviews in the field with notebook computers, has been one of DMQC's most innovative accomplishments. This computer-assisted personal interviewing (CAPI) system is integrated into the overall DMS and offers many advantages over traditional paper-and-pencil interviews. All measures for the cohort assessment have been computerized using an application called FormBuilder. With FormBuilder, questions, answer options, and skip-pattern codes are entered to create computer "forms." Because FormBuilder was designed to be used by nonprogrammers, it is relatively easy to use. To build a form, the user enters the text of the question along with a question name, question label, and response format. The response format can be any one of the following: character, numeric, date, time, multiple choice, or text only. Once a response format has been selected, FormBuilder prompts the user for additional information, such as field length. In the case of multiple choice formats, the user first builds a "library" of multiple choice response tables. Then, as each question is entered, the system directs the user to the appropriate screen to select the correct multiple choice table from this "library." In addition to entering the questions and associated information, the user creates the skip-pattern coding based on the paper form, using the "route" option in the design menu. Once this information has been edited, the user can build the form. Data quality is improved with this interview method because the computer prevents certain types of errors from being made during the interview. For example, skip patterns are automatic, so the interviewer can neither erroneously skip questions that should be asked or ask a question that is not applicable. In addition, some information is automatically recorded, such as "time started" and "time stopped." The use of CAPI also eliminates the need for a separate data-entry operation as computerized interview data are uploaded into the data base, saving time and money and enhancing data quality by removing data entry as a source of error.

The Data Collection System (DCS) is the central control system for the notebook computer, which enables the research assistant to administer the interview. It is written in DOS to lessen the memory and space requirements needed on notebook computers. This application consists of DCS itself, which manages the overall data collection, and FormRunner, which executes the individual questionnaires. Working together, DCS and FormRunner provide the flexibility required for a variety of interview situations. For example, using the "suspend" feature, the interviewer can stop in the middle of a particular questionnaire and return easily to that exact point later. This capability will be useful in situations where the study participant does not have enough time to complete the entire assessment in one sitting. The "refuse measure" feature allows the interviewer to mark the remainder of a questionnaire as refused by the participant and proceed to the next instrument. The "later" feature gives the interviewer the option of skipping certain measures and returning to them at the end of the interview. This option is valuable because sensitive measures can only be administered when

no other household member is in the room. Thus, both DCS and FormRunner were designed to handle the complexities of the multimeasure data being collected.

In addition, DCS includes several safeguards to protect the data during collection. First, all personal information on the notebook computers is encrypted and the notebook computer itself is password protected. During data collection, each response is saved as the interview takes place. This means that in the event the interviewer has to conclude the interview abruptly or the computer loses power, all responses recorded to that point are secure. Finally, the interviewer always is prompted to back up the data to floppy disk when the session is terminated. Short of turning off the computer in the middle of the interview, there is no way to exit the system without being instructed to back up the data.

DMS and DMS_Link

In addition to the FormBuilder and DCS applications, two other elements are critical to the Project's data management: (1) DMS, which manages the data in the central office, and (2) DMS_Link, which provides the interface between the data collection and data management systems.

DMS. DMS is a large and complex system that includes the overall relational data base structure in which the Project data and the set of integrated applications used to manage the data are maintained. A relational data base design was chosen to facilitate linking records across different levels of aggregating and integrating new data elements without costly and time-consuming rewrites of the existing data base. Although many of the components in the overall DMS design are complete and functional, new elements are still being added. As new elements are completed, they are quickly brought online for use. The primary features of DMS are:

- Storing and safeguarding all project data.
- Tracking all case activity, including research assistant assignment, contact information, and tracking and retention activities.
- Maintaining current and historical personal information—name, address, and phone number—for study participants.
- Generating reports for all project units.

Security systems have also been developed by DMQC to protect both Project data and the confidentiality of study participants. Project data are protected from loss by a system of backups. The data base file server is equipped with a RAID (redundant array of inexpensive disks) to guard against data loss in the event of a hard-drive failure. This technology protects the data by spreading the data across multiple drives ("striping") and by storing the same data on a second drive ("mirroring"). In addition, all project data are backed up to tape on a

daily basis and maintained for 2 weeks, and backup tapes are moved to offsite storage once a week. The confidentiality of the study participants is protected by security features at both the network and program levels. The network operating system, Novell, allows the network administrator to control directory and program access. DMS also has a built-in security system to provide further access control to data files. For example, within DMS, only a limited number of people who have "super-user" status are allowed to read all of the data files.

DMS also is used as a tool to manage field activity and to report on progress in the field. Field management staff can record which cases have been assigned to which research assistants and the current status of each case. This information is used to generate reports to monitor the field activity progress. This DMS feature is currently in operation for the Cohort Assessment Unit (CAU) and will soon be used in the Infant Assessment Unit (IAU) as well. In the interim, IAU is tracking their case activity with a stand-alone application developed by DMQC staff. Data entered into this temporary system are fully compatible with the DMS design and will be moved easily into DMS when that part of the system is complete.

Finally, DMS is used to help find study participants who move or are hard to reach. Personal information such as names and addresses can be updated, but the old information is retained. Each update will include the new information along with the date of entry and the source of that information. During each wave of data collection, research assistants also will obtain the names and addresses of people who will always know how to contact study participants. Maintenance of these tables will be critical to sample-retention efforts.

DMS_Link. DMS_Link is an application that permits the interface between DCS on the notebook computers and DMS on the central office computer. Prior to an interview, research assistants will download study participant information and the correct protocol from DMS to their notebook computer. Upon returning from a completed interview, they will upload the interview data back to DMS via DMS_Link. The creation of this interface was a challenging task for DMQC because the notebook computers run under DOS-based programs and the network runs under Windows. Additional programming was required to bridge the difference in operating systems. Another design challenge in DMS_Link was to design a way to handle temporary storage of study data in situations where an interview requires more than one visit. DMS_Link is now well equipped to handle a variety of scenarios involving partially completed assessments.

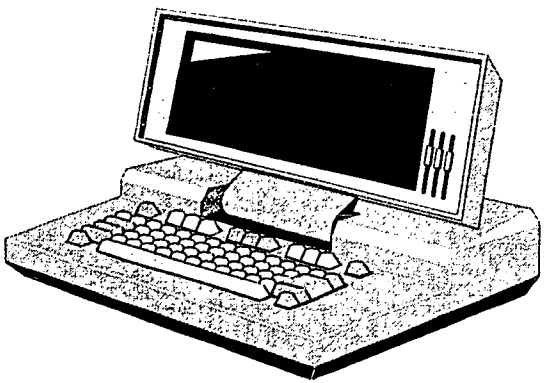
See figure 13.1 for an illustration of the Data Management System and the Data Management System_Link.

Figure 13.1

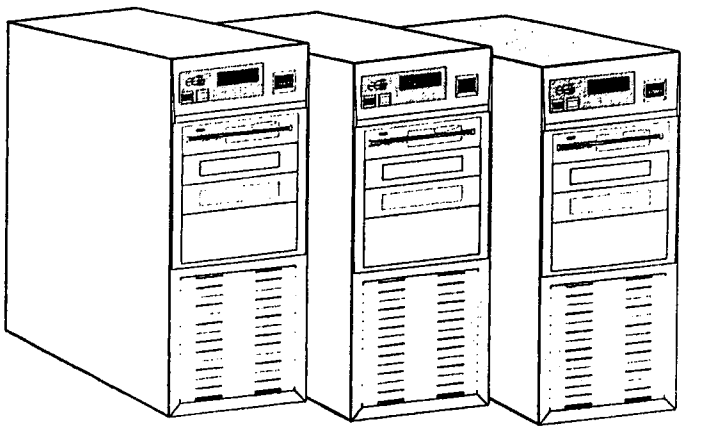
Integrated Data Collection/Management System

Forms and study participant information
(e.g., name, address, ID number) downloaded

DCS
manages notebook
data collection



FormBuilder
computerizes
the measures



FormRunner
executes the
measures

DMS
manages all
cohort data

Assessment data
checked and uploaded

Part Five: Project Relations, Communications, and Future Research Directions

Chapter 14. Archives

by Susannah Bates, M.Ed.

An archive of Project documents, dating back to 1987, is being maintained. Files and binders—including information such as core group/scientific directors' meeting materials and minutes, status reports, subcontract information, publications, grants, and correspondence—are part of a growing collection that documents the Project's history. All files and binders are being organized, labeled, assigned an ID number, and recorded into a data base. A data base "directory" of Project documents has been created to facilitate quick review of contents and retrieval of information. In addition, special efforts are being made to capture important events in the Project's history, such as public controversies covered by the media. During 1996, plans are under way to incorporate the Project's earliest documents, dating as far back as 1984–1985; this will involve interviewing some of the Project originators, including Michael Tonry, Richard Linster, James Q. Wilson, Norval Morris, and Lee Robins. The Project archives will be maintained throughout the course of the study.

Table 14.1 is a sample of page printouts from the archives data base. Each item (file or volume) listed has been assigned an ID number and a letter (either G, E, or V) indicating the "type" of file: (1) General, (2) Event, or (3) Volume. This data base has become a useful tool for locating Project documents.

(1) General: These files focus on ongoing subjects or aspects of the study, such as general administration, agency records and sample retention, subcontracts, and newsletters, that are not specific to a particular year and are therefore not dated. Information will be added to these files in chronologic order throughout the course of the study. ID numbers for all general files start with the number 1.

(2) Events: These files focus on particular events that took place (meetings, conferences, etc.) and are in chronological order dating back to 1988. ID numbers for all events files are based on the year the event took place. For example, event file 9006 contains the minutes from the 1990 Workshop on Statistical Analysis; event file 9101 contains the agenda for a 1991 advisory group meeting.

(3) Volumes: Some event information (such as meeting materials) was originally organized into volumes and is stored in its own archives category. These volumes also include information such as grant proposals and collections of articles and publications. Reports from Phase I working groups also are catalogued in this section. ID numbers for most volumes begin with the year the event took place; a few volumes have ID numbers starting with number 1 if the information in that volume is not specific to a particular year.

Table 14.1

Archives

3/6/96

ID#	Type	File/Doc Title	Date	Category
1100	G	Jan Dunham		Data Management & Quality Control
1101	G	Intelligence Plan		Data Management & Qualtiy Control
1110	G	Steve Raudenbush		Data Analysis
1111	G	Richard Congdon		Data Analysis
1120	G	Correspondence		Project Relations
1121	G	Patricia Lau		Project Relations
1122	G	Brochures		Project Relations
1123	G	Lillian Calhoun		Project Relations
1124	G	Marilyn Katz		Project Relations
1125	G	Newletters (Chicago)		Project Relations
1126	G	Bad Press		Project Relations
1127	G	Good Press		Project Relations
1130	G	Dan Kindlon		Infant Assessment Unit
1131	G	Ann Strohm		Infant Assessment Unit
1132	G	Intelligence Plan		Infant Assessment Unit
1140	G	Individual Measures		Research Measures
1141	G	Family Measures		Research Measures
1142	G	School Measures		Research Measures
1143	G	Peer Measures		Research Measures
1144	G	Neighborhood Measures		Research Measures
1145	G	Other		Research Measures
1150	G	Correspondence		MacArthur Foundation
1151	G	Judge Baker Annual Reports		MacArthur Foundation
1152	G	PHDCN Annual Reports		MacArthur Foundation
1153	G	Miscellaneous		MacArthur Foundation
1160	G	Correspondence		NIJ
1161	G	Quarterly Reports		NIJ
1162	G	Miscellaneous		NIJ
1168	G	RWJ Correspondence		RWJ
1170	G	Abt Documents		Subcontract
1171	G	Abt Correspondence		Subcontract
1172	G	Abt - Revised Subcontract	5/1/95	Subcontract
1173	G	Meeting with Abt	3/17/95	Subcontract
1174	G	Abt- Weekly Meeting Minutes		Subcontract
1175	G	Abt - Monthly Reports		Subcontract
1176	G	Abt- Data Mgmt & Quality Co		Subcontract
1180	G	Duke Univ. (Angold)		Subcontract
1181	G	Kandel		Subcontract
1182	G	Lessler		Subcontract
1183	G	Michigan State (Raudenbush)		Subcontract
1184	G	NORC -Key Informant Survey		Subcontract
1185	G	NORC -Syst Social Observ		Subcontract
1186	G	MCIC		Subcontract
1187	E	Washington Univ (R. Price)		Subcontract
1188	E	Univ. of P'burg (R. Loeber)		Subcontract
1189	E	Duke Univ. (J. Coie)		Subcontract
1190	E	Univ. of Montreal (Tremblay)		Subcontract
1195	G	Predoctoral Students		Student Fellows
1196	G	Postdoctoral Students		Student Fellows
8701	V	Grant Proposal	8/14/87	Proposal to NIJ & MacArthur
8801	E	Minutes	5/12/88	Pathways Working Group
8802	E	Transcript	8/5/88	Pathways Working Group
8803	E	Notes	8/5/88	Pathways Working Group
8806	V	Final Report	12/1/88	Onset Working Group
8807	V	Meeting Materials	8/5/88	Pathways Working Group
8808	V	Meeting Materials		Materials for several planning meeti

Table 14.1 continued
Archives

3/6/96

ID#	Type	File/Doc Title	Date	Category
8809	V	Meeting Materials		Materials/Notes from Meetings in D
8810	V	Meeting Materials		Desistance Working Group
8811	V	Meeting Materials	11/11/88	Desistance Working Group
8812	V	Meeting Materials	8/5/88	Pathways Working Group
8813	V	Meeting Materials	9/22/88	Research Advisory Board Meeting
8814	V	Meeting Materials		Onset Working Group
8815	V	Meeting Materials		Onset Working Group
8816	V	Paper		Explaining Delinqency & Drug Use
8817	V	Final Report	12/15/88	Desistance Working Group
8818	V	Final Report	12/15/88	Pathways Working Group
8819	V	Final Report	12/15/88	Onset Working Group
8901	V	Papers/Articles	3/1/89	Research Design Working Group
8902	V	Meeting Materials	4/20/89	Research Design Working Group
8903	V	Meeting Materials	5/15/89	Research Design Working Group
8904	V	Meeting Materials	8/17/89	Research Design Working Group
8905	E	Meeting Materials	10/3/89	Graylyn Conference on Longit Res
8906	V	Report	6/8/89	Interim Report to Advisory Group
8907	V	NLSY Child Handbook	1/18/89	Guide to Use of Nat'l Longit Surv of
8908	V	Report	5/29/89	Violent Recidivism ...
8909	V	Report	6/8/89	Violent Recidivism...
8910	V	Proposal	5/8/89	Proposals for Side Studies
8911	V	Meeting Materials	6/8/89	Research Advisory Board Meeting
8912	V	Meeting Materials	5/26/89	Research Organization Working Gro
8913	V	Meeting Materisl	6/27/89	Research Organization Working Gro
8914	V	Reports	8/25/89	Critiques on Reports
8915	V	Final Report	12/15/89	Report on Phase II Activities
8916	V	Letters/Articles		Requests for Info for PHDCB - 89-9
9001	E	Workshop Summary	6/11/90	Gender Issues Workshop
9002	E	Notes	6/11/90	Gender Issues Workshop
9003	E	Workshop Report	6/11/90	Gender Issues Workshop
9004	E	Minutes	7/15/90	Phase III Working Group
9005	E	Workshop materials	7/15/90	Phase III Working Group
9006	E	Minutes	7/16/90	Workshop on Statistical Analysis
9007	E	Notes	7/16/90	Workshop on Statistical Analysis
9008	E	Workshop Materials	7/16/90	Workshop on Statistical Analysis
9009	E	Minutes	8/14/90	Statistical Issues Workshop
9010	E	Workshop Materials	8/14/90	Statistical Issues Workshop
9011	E	Transcript	8/14/90	Statistical Issues Workshop
9012	E	Workshop Report	9/14/90	Biological Measures Workshop
9013	E	Notes/Articles	9/14/90	Biological Measures Workshop
9014	E	Workshop Materials	9/14/90	Biological Measures Workshop
9015	E	Transcript	9/14/90	Biological Measures Workshop
9016	E	Transcript	9/15/90	Core Group Meeting
9017	E	Workshop Report	11/2/90	Community Measures Workshop
9018	E	Workshop Materials	11/2/90	Community Measures Workshop
9019	E	Transcript	11/2/90	Community Measures Workshop
9020	E	Workshop Materials	12/14/90	Fam Struct & Ethn Workshop
9021	E	Workshop Report	12/14/90	Fam Struct & Ethn Workshop
9022	E	Agenda	12/16/90	Core Group Meeting
9023	V	Manuscript	4/12/90	Manuscript for PHDCB articles
9024	V	Meeting Materials	6/11/90	Gender Issues & Antisocial Behavior
9025	V	Report	10/1/90	The Willie M. Program
9026	V	Meeting Materials	11/3/90	Core Group Meeting
9027	V	Meeting Materials	12/16/90	Core Group Meeting
9101	E	Agenda	1/31/91	Advisory Group Meeting
9102	E	Minutes	1/31/91	Advisory Group Meeting

Chapter 15. Project Relations Strategy

by Patricia Lau, M.P.A.

A community relations strategy was developed in 1995 to build local support and secure the cooperation from study participants that is essential for the success of the Project. This plan has focused on almost daily meetings with community leaders in politics, education, religion, law enforcement, business, and social service organizations in Chicago to clarify the Project's purpose to the public and has established a collaborative relationship between these leaders and the Project. In a city where illegal immigration is prevalent, where suspicious and fearful communities have vigilant "watch" groups, and where gang turfs need to be acknowledged, the process of recruiting and retaining subjects for the study is complex. Added to this picture are the inaccurate and potentially damaging charges from the Chicago Coalition Against the Violence Initiative, which have required the Project's energies to counteract. The Project worked with Media Strategies (a public relations firm) during the past year to train senior Project staff to deal with the media, especially on the national level, and to train staff to answer questions about the Project in the field.

The Project Relations Unit provides a vital communication link between the Project and the community through the following activities:

- Educating and informing the public about the purpose and goals of the study, including designing and developing brochures, handouts, and presentation materials that explain the Project's mission to diverse audiences.

- Developing and nurturing community-based contacts within various neighborhoods in an effort to marshal support and interest in the Project, and initiating and maintaining relationships with community leaders and policymakers as a foundation for continuous and fluid communication throughout the course of the study.

- Addressing questions and concerns from special interest groups that may be misinformed about the Project's mission and goals.

- Publishing a newsletter to provide study participants, community leaders, and policymakers with current information about the status, findings, and results from the study.

- Facilitating an internal communications network among the Project's senior management staff, university public affairs offices for the scientific directors, the MacArthur Foundation, and the National Institute of Justice.

- Responding to all inquiries from the media at the national, city, and community levels, reviewing and evaluating all interview requests in conjunction with scientific directors and media consultants, and coordinating interviews as needed.

Chapter 16. Publications and Communications

by Patricia Lau, M.P.A., and Susannah Bates, M.Ed.

The Project has developed a variety of communications products to provide information about the Project as it progresses.

Within the Project Relations Unit, a variety of documents have been either created or revised within the past year to provide relevant, accurate, and timely information about the Project to the local public. These materials include a question-and-answer handout; an overview of the study; a brochure titled *Intricate Pathways*; an essay written by Felton Earls, titled *Connecting Social Science to the World*, which was published as an additional brochure in spring 1995; and recent newspaper articles from local and national newspapers to provide different perspectives on the study. These materials often are combined in a presentation folder or press kit and distributed to key community leaders and media representatives.

The Project Relations Unit also has developed "community-friendly" literature that features a one-page fact sheet, a "backgrounder," and a brochure translated into Spanish, Polish, Russian, Chinese, and Cambodian. These materials help build and secure cooperation among neighborhood residents and facilitate study enrollment. Recent newspaper clippings about the Project also are provided to help explain the Project's goals. In addition, letters that have been translated into Spanish and Polish and include endorsements from national, State, and local elected leaders are distributed to help gain residents' interest and support.

The Project has developed several ways to communicate its progress and findings to the public. Two quarterly newsletters have been initiated, with a distribution of over 1,000 each. *City News: Chicago in Ten Years*, produced by the Project Relations Unit, keeps community leaders and subjects informed about the Project. This publication includes a change of address card to keep track of participants when they relocate. *The Chicago Project News*, edited by Susannah Bates in the Boston office, keeps public health and criminal justice professionals up to date on the Project's issues and findings.

The Project also has designed a home page on the World Wide Web that can be accessed with a Web browser such as Netscape or Mosaic at: http://phdcn.harvard.edu. This home page, which is updated regularly, includes an overview of the Project and newsletters and is available to the public. Information such as scientific directors' meetings, weekly meetings from both the Boston and Chicago offices, and Project staff members' calendars are posted on the home page with password protection for Project staff only and are distributed to Project staff by e-mail as well.

Chapter 17. Future Research Directions

by Felton Earls, M.D.

Some future directions for the Project include the challenge of integrating physiological measures of aggression, such as heart rate, lead levels, and hormonal influences, into the study. The Project is currently conducting a study of lead levels in children through the Chicago Department of Health and, as described in the Exposure to Violence chapter of this report, will be collecting saliva samples from children to study cortisol and its relationship to stress and aggressive behavior. A series of three workshops by Felton Earls and Bruce McEwen on neurobiological research to explore the interaction between environment and neural mechanisms took place at Rockefeller University in early 1996. This work has included an investigation of the influence of testosterone on violent behavior that found no differences in hormone levels between a group of psychiatrically hospitalized boys under age 9 and controls (Constantino, Grosz, Saenger, Chandler, Nandi, and Earls, 1993), as well as more recent studies on patterns of cortisol secretion in infants institutionalized since birth (Carlson, Dragomir, Earls, Farrel, Macovei, Nystrom, and Sparling, 1995). In both instances, the Project is investigating extreme groups, following the logic that if positive links are not revealed within these groups, there is little merit in addressing such questions in a general population sample.

A continuing concern of the Project has been how to find sensitive approaches to measure family influences on child social and psychological development. Although not covered in this report, Project scientists engaged in several systematic reviews of the existing literature (McGuire and Earls, 1993) and field-tested measures that appeared particularly promising (Barnes McGuire and Earls, 1994; McGuire and Earls, 1995).

A domain of psychological development that needs refinement prior to its incorporation into the Project's protocols has been referred to as perceived competence, self-efficacy, or human agency by various investigators. The Project is profiting from the work of an extensive review of theories and measures in this area (Haidt and Rodin, 1995) and is currently developing a new measure to test (much as was done with the Exposure to Violence measure).

The Project also has investigated the methodology for assessing the extent of father involvement in infants' lives and the relationship between father involvement and cognitive outcome (Yogman, 1984; Yogman, Kindlon, and Earls, 1995).

Finally, the Project plans to incorporate ethnographic studies into the Project to capture the social life of a subgroup of the 80 study neighborhoods beyond the capacity of survey approaches. Some observations will include:

- Who goes to work?
- What do children do?
- How openly do drug markets function?
- How visible are police?
- What is the availability and use of public recreation space and facilities for youths?
- How much adult supervision do children receive?

With assistance from youths and parents as consultants and informants, the Project will gain additional information about what it is like to grow up in Chicago's communities and what it is like to raise children and teenagers given the strains of life in a large city.

References

Barnes McGuire J and Earls F (1994). Research note: The test-retest stability of the five minute speech sample in parents of disadvantaged, minority children. *Journal of Child Psychology and Psychiatry*, 35, 5, 971-979.

Carlson M, Dragomir C, Earls F, Farrel M, Macovei O, Nystrom P, and Sparling J (1995). Effects of social deprivation on cortisol regulation in institutionalized Romanian infants. *Society for Neuroscience Abstracts* (218.12).

Constantino JN, Grosz D, Saenger P, Chandler DW, Nandi R, and Earls F (1993). Testosterone and aggression in children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 32, 6, 1217-1222.

Haidt J and Rodin J (1995). Control and efficacy: An integrative review. A report to the John D. and Catherine T. MacArthur Foundation, Program on Mental Health and Human Development, University of Pennsylvania (unpublished report).

McGuire J and Earls F (1993). Exploring the reliability of measures of family relations, parental attitudes, and parent-child relations in a disadvantaged minority population. *Journal of Marriage and the Family*, 55, 1042-1046.

McGuire J and Earls F (1995). Coercive family process and delinquency: Some methodological considerations. In: McCord J. (ed.), *Coercion and Punishment in Long-term Perspectives*. Cambridge University Press, pp. 348-361.

Yogman MW (1984). The father's role with preterm and full-term infants. In: Galenson E and Tyson R (eds.), *Frontiers of Infant Psychiatry* (vol. 2). New York: Basic Books.

Yogman MW, Kindlon DJ, and Earls F (1995). Father involvement and cognitive/behavioral outcome of preterm infants. *Journal of the American Academy of Child and Adolescent Psychiatry*, 34, 58-65.

For more information on the National Institute of Justice, please contact:

National Criminal Justice Reference Service
Box 6000
Rockville, MD 20849–6000
800–851–3420
E-mail: askncjrs@ncjrs.org

You can view or obtain an electronic version of this document from the NCJRS Bulletin Board System (BBS) or the NCJRS Justice Information Center World Wide Web site.

To access the BBS, direct dial through your computer modem: 301–738–8895
(modems should be set at 9600 baud and 8–N–1),
or telnet to bbs.ncjrs.org or gopher to ncjrs.org:71

To access the World Wide Web site, go to http://www.ncjrs.org

If you have any questions, call or e-mail NCJRS.

U.S. Department of Justice
Office of Justice Programs
National Institute of Justice

Washington, D.C. 20531

Official Business
Penalty for Private Use $300

BULK RATE
POSTAGE & FEES PAID
DOJ/NIJ
Permit No. G–91