MF-1

social science
research institute

USC

LONGITUDINAL RESEARCH IN THE UNITED STATES:

RELEVANCE TO PRIMARY PREVENTION OF DELINQUENCY

BY

SARNOFF A. MEDNICK

AND

MICHELE HARWAY

VOLUME I

LONGITUDINAL RESEARCH METHODS

UNIVERSITY OF SOUTHERN CALIFORNIA

83031-1
83033/3

Final Report

Submitted to:

The National Institute of
Law Enforcement and Criminal Justice

Department of Justice
633 Indiana N. W.
Washington, D. C.   20531

by

Sarnoff A. Mednick

and

Michele Harway

University of Southern California
Social Science Research Institute

December, 1981

NCJRS

FEB 10 1982

ACQUISITIONS

TABLE OF CONTENTS

TABLE OF CONTENTS

Acknowledgements

Chapter 1

Rationale for the Study

Michele Harway, Ph.D.

Sarnoff A. Mednick, Ph.D., Dr. Med.

University of Southern California
Social Science Research Institute

The Prevention of Criminality

Society devotes large amounts of energy and funds to the maintenance of the apprehended offender and to attempts at rehabilitation. Considerable ingenuity is lavished on research on relatively esoteric and technically sophisticated treatments. New therapies from psychology, psychiatry and sociology are inevitably transplanted and tested on delinquent and criminal populations. All of these methods have in common that they wait for the appearance of criminality before acting. In order to reduce human suffering and the increasingly immense burden to the community, it is critical that some proportion of society's resources be devoted to finding ways to prevent the initial onset of the antisocial behavior.

Primary prevention centers on the early identification of those in whom some interaction of biology, psychology and social conditions produce a predisposition to later antisocial behavior. Those identified as being predelinquent become candidates for exploratory preventive efforts. It is likely that this approach will prove most meaningful for the small group of offenders who produce the majority of criminal acts and the more serious criminal acts (Wolfgang, Sellin and Figlio, 1972).

Early detection. It is unlikely that criminality will yield to blanket interventions applied to entire populations (such as floridation of water). Also, since relatively small fractions of the population become antisocial, it would be wasteful and unfair to subject those not at risk to interventions. In addition, the interventive procedures may actually entail risks and be psychologically intrusive and time consuming.

For these reasons interventions should be restricted to individuals at high risk for evidencing antisocial behavior. As a consequence, a first step in the development of primary prevention methods must be to devise assessment procedures which will effectively select the future antisocial individual (It should also be pointed out that if such early distinguishing characteristics of the future criminal can be identified, these may help suggest intervention procedures.) Prospective longitudinal research can make a valid contribution in this process of devising early detection procedures.

Problems in research. How do we go about devising procedures for the early detection of the future antisocial individual? Almost all the research which exists in the literature on criminality begins with the apprehended or incarcerated delinquent criminal. It may be difficult, however, to identify precursors of antisocial behavior through studies carried out with individuals who have lived through the process of becoming and being antisocial. The behavior of these individuals may be markedly altered in response to the correlates of the criminal life, such as educational, economic and social failure, drug and alcohol ingestion, institutionalization, and anomie and alienation from society. In research with non-criminals these same factors have been shown to

measurably affect research results. If we could find control groups which were equated with criminal groups for all of these correlates of antisocial behavior, then observed differences could conceivably be useful in the early detection of criminality (barring maturation factors). But in reality such control groups are not readily available. Consequently, when we compare criminals and law-abiding citizens it is often difficult to judge what portion of the reported differences have unique relevance to and could forecast criminality and which may be ascribed to the consequences of being a criminal.

A good example of this type of confounding is demonstrated in a study by Silverman, Berg and Kantor (1966). In 1964 Silverman had observed acute and chronic schizophrenics on measures of perceptual scanning and concept formation. The acute and chronic schizophrenic groups differed widely, however, in their duration of hospital stay. Concern with this factor, and a desire to evaluate the possible effects of institutionalization on the obtained relationship, led Silverman, Berg and Kantor (1966) to take the unique step of repeating the same tests with 50 long-term and 50 short-term non-psychiatric prisoners at San Quentin. All the prisoners were felons relatively free of psychiatric problems. The acute schizophrenics were matched with the long-term prisoners, and the chronic schizophrenics were matched with the long-term prisoners for length of institut-ionalization. They found that the differences observed on these conceptual and perceptual measures between the short-term and long-term San Quentin inmates were almost precisely the same as those differences observed between the acute and chronic schizophrenics. The mean scores for the imprisoned and the hospitalized were almost identical. "The failure to observe noteworthy differences between them (the schizophrenics and the prisoners) indicated that exaggerated perfor-mances on these procedures are not pathognomic of particular types of schizophrenia... comparable response patterns may be found among non-schizophrenics who live in a similar environment" (p. 656). Psychological measures of perception and con-cept formation may be influenced in important ways by prison institutionalization. If we try to find methods of early detection by examining criminals who have been exposed to prisoner experiences, differences between them and controls may (at least in part) be a function of the institutionalization that they have suffered.

Studies of the biochemistry and physiology of criminals which make use of populations who are incarcerated or have been incarcerated may find their results heavily influenced by the exercise regimens, the special nutritional pro-grams, and the stresses and strains of life in prison institutions. But prison is only one type of deviant influence which some antisocial individuals experience. As indicated above, drug and alcohol ingestion, economic deprivation, social skewness, disturbed family life and educational disadvantage are only some of the other correlates of delinquency and criminality which could produce differences between antisocial individuals and controls. These differences are, in the main, consequences of the life of the antisocial individual and could not be used as predictors of future criminal behavior.

Reports from interviews with families of antisocial individuals might also be biased by the fact that their child is an acknowledged criminal. Families may be either defensive or belligerent, but in any case may testify in an altered manner from that which they would have shown before the child evidenced criminality. If we observe disturbances in the family relationships or even breakup of the family, we might be tempted to assert that this is a possible cause of the

antisocial behavior; but we should also be alert to the possibility that it is the delinquent behavior of the offspring which has been responsible for the disturbance in the family.

Need for prospective longitudinal research. These considerations suggest that it would be unwise to attempt to search for early detection characteristics by examination of individuals already involved in a problem lifestyle. The criminal is so changed by the criminal lifestyle--the drugs, the alcohol, contact with the criminal justice system, the brain damage potentially suffered in violent espiodes--that some portion of the differences between criminals and controls may simply reflect the consequences of the criminal life rather than antecedents. Therefore, we must intensively assess the criminal before he or she becomes exposed to these criminal life variables. This means examining a young non-criminal population intensively and following this population until one can register who among them have become criminal. Then one can go back to the original intensive assessment and see what early life characteristics distinguished the criminal from law-abiding co-subjects.

For a more concrete example, if we wished to develop methods of early detection of the recidivistic criminal, we might first intensively examine a large population of school age children. By following this population for a suitable period of time and registering the apprehended offenses as they occur, we could identify the recidivists among them. We would then be able to return to the initial intensive assessment and determine what combination of premorbid characteristics distinguished the recidivist. This combination of characteristics (if it proves replicable) could then be used to select those likely to be future recidivists from a new population of school age children. These "likely" future recidivists could then be subjected to intervention procedures aimed at reducing the probability of recidivism. As mentioned above, the characteristics which distinguish these individuals might produce hypotheses regarding methods of intervention.

Although this procedure can be outlined in one paragraph, it represents a long, difficult and very expensive process.

Utilization of existing longitudinal studies.

While it is ultimately necessary to initiate some longitudinal research such as that sketched above, there are prudent steps which should be taken before launching such massive undertakings. In the United States there are a large number of existing longitudinal projects which have been active for many years. Some of these have involved extremely intensive long-term assessments of the perinatal, cognitive, personality, family relations and even physiological status of their subjects. Subjects are frequently very well identified and followed closely by the projects. In many instances a researcher could determine the criminal behavior of such a population and then be in a position to describe the childhood and youth characteristics of these offenders. Note that it is not necessary that the original longitudinal researchers have any notion that their work could be used in criminological research. What is needed is an identified population of some considerable size which was intensively examined with relevant measures at some early point in their lives.

We will relate one example of how this kind of utilization of an existing longitudinal project has been successfully attempted. In 1946 James W. B. Douglas instituted a massive birth cohort study including 5,362 children born in the British Isles between the 3rd and 9th of March, 1946. Over the course of their development this cohort has yielded a rich harvest of information which has been of critical importance to the shaping of English policy in education and social welfare (Douglas, 1975).

As might be expected, the cohort's progress through elementary and high school, as well as occupational selection, has been carefully monitored. More recently, sociologist Michael Wadsworth (1975) described the patterns of delinquency of the males in the population. Taking advantage of the project's data file, Wadsworth noted that those who became delinquent were overly heavily represented among individuals who had suffered early emotional deprivation. Stimulated by a world literature suggesting an association of autonomic nervous system "sluggishness" and antisocial behavior, Wadsworth (1976) also found a strong relationship between a pulse rate taken by school physicians during a health examination at 11 years of age and serious delinquency assessed a decade later. These two results must be viewed as being of extreme importance for criminology, both because the predictor-criterion time span of a decade is impressive, and because the finding is produced in the context of a total birth cohort . As illustrated by the Wadsworth study, older longitudinal projects could be utilized to explore questions which might inform the design of new longitudinal undertakings.

Wadsworth's example suggests that it may be fruitful to identify existing longitudinal projects which are well-suited for secondary analyses. One of the goals of the study described here was the identification of such projects.

Methods of Identification of U.S. Longitudinal Research

The study which we have just completed involved a survey of existing longitudinal research projects, and a compilation of U.S. longitudinal research which lends itself to the investigation of the antecedent factors in delinquency, criminality and recidivism.   Such a compilation is supplemented by commentary about how the studies might be utilized in criminogenic research. The compilation and commentary are presented in the following chapters. The process we utilized for the study is delineated here.

Procedure. The first problem we faced was delineation of criteria for the inclusion of projects. After consideration of the different types of longitudinal research and their advantages and disadvantages, a taxonomy of projects was developed. The taxonomy, which also formed the basis for a table of contents for this volume, is provided in Table 1. That is, the taxonomy is organized to accomodate the longitudinal projects included in this volume.

Table 1

Descriptions of Longitudinal Research

A. Correlative Longitudinal Research

  1. Normal, Representative Populations

    a. Birth Cohorts

    b. School-aged and College-aged Cohorts

    c. Military Cohorts

    d. Adult Cohorts

    e. Community Cohorts

  2. Non-representative Populations

    a. Normal Cohorts

      i. Birth Cohorts

     ii. Pre-school Cohorts

    iii. School-aged and College-aged Cohorts

     iv. Adult Cohorts

     v. Aged Cohorts

     vi. Marriage and Family Cohorts

    b. Specialized Cohorts

      i. Perinatal Damage

     ii. Mental Patients

    iii. Delinquents and Criminals

     iv. Children at Risk

     v. Disadvantaged Students

     vi. Special Children

B. Experimental- Manipulative Research

Description of Taxonomy. Our intention was to include as broad a range of longitudinal projects as possible. Therefore we decided to define longitudinal research simply as requiring some assessment of subjects from a defined population at a minimum of two points in their lives.

We included follow-up studies, epidemiological and retrospective studies, as well as other more traditional longitudinal studies. We also included projects whether they were ongoing, completed or just beginning. The taxonomy displayed in Table 1 is organized primarily according to characteristics of the sample studied rather than by methodological approach except for the two large categories of Correlative and Experimental-Manipulative Research.

Correlative Longitudinal Research includes studies where outcome characteristics of the group under study are correlated to previously measured antecedent factors. This contrasts with Experimental-Manipulative Research which involves the implementation of an intervention for one part of a population with the remainder of the population forming the control group.

Under the two broad methodological headings are categorizations of types of cohorts as described below: Normal representative populations. The normal representative populations mentioned in the table tend to be large (in the thousands), or to be randomly selected samples from a larger population. The birth cohorts are defined by all births in an area within a given time period (typically a year). Birth cohorts very often include all births at large obstetrical departments for a period of some years. School-aged or college-aged cohorts usually represent all students in an existing class, for example, all kindergarteners represent all students in an existing class, for example, all kindergarteners in a county during a given year, or all college freshmen in a particular region. Military cohorts might include a study of air pilots during a specified era. Adult cohorts would include large unspecified groups of adults such as a random sample of residents of the state of Nebraska. Community cohorts will take all individuals of all ages living in a given moderate size community.

Non-representative populations include samples selected in a non-random fashion from a larger population. Thus, adults who volunteer for participation in a health study would comprise a non-representative population. This category is further subdivided into normal and specialized cohorts. Normal cohorts can be defined by their ages, as were the representative cohorts. Specialized cohorts are those groups which are selected for study because of some specific characteristic of interest. Studies of perinatal damage cohorts may be of some special interest since such early possible causes of brain damage might very well be related to later impulsive antisocial behavior. Following up a population with pregnancy difficulties, abnormal delivery conditions or neonatal neurological difficulties permits one to assess the criminologic consequences of these early disturbances. Groups of mental patients can be studied to determine the relationship of mental illness to antisocial behavior.

Long-term study of delinquents and criminals can provide information on the probabilities of recidivism. If the clients are intensively examined early in their criminal career, then the relationship between these early signs and followup status can suggest signs predictive of recidivism.

Of particular importance in this group of specialized cohorts is the group of children at risk. Such children might be selected to be at risk because of antisocial behavior in their parents, because they belong to social groups which have a high probability of criminal or delinquent acts or because of other factors which predict to their later criminality. A great advantage of the risk design over the study of normal, representative populations is the higher yield of antisocial individuals in the study. This means that in order to eventually obtain a sample of criminals of a given size, one can begin with a smaller total cohort. This can be of great importance if time-consuming assessments are envisioned. Disadvantaged students would represent one group of children who are seen by researchers as being particularly at risk for a variety of unfortunate reasons such as poverty, unemployment and lower educational attainment. Special children (a term used in education) are selected for study because they are gifted or because they are hyperkinetic.

Examples of these types of longitudinal research may be found in most of the chapter of this report.

Contacting the projects. Subsequent to the development of the inclusion criteria described with the above taxonomy, we set out to contact projects. Since our focus was inter-disciplinary, we published notices soliciting information about longitudinal research in approximately 20 journals and newsletters in the fields of psychology, sociology, education, criminology, medicine, social welfare, psychiatry and public health. We also contacted professional organizations. We sent posters to all of the departments in major medical schools, to departments of psychology, sociology, education, criminology, social welfare, and public health in 100 universities. We also wrote personal letters to 396 invidivuals involved in longitudinal research. Finally, we conducted eight computer searchers to identify completed and ongoing government-funded projects and to identify any projects which might not otherwise have come to our attention.

Our next step was to identify projects which we considered as being of significance for studying possible etiological variables in criminology. The criteria considered in making this judgment were: size, age of the cohort studied, breadth of information available, possibility for follow-up and type of variables in the initial assessment.

We next visited a number of projects which had been identified in the previous step. The purpose of these visits was to get a better first-hand familiarity with the possibility of doing criminological research with each data file. We were particularly interested in materials which might not be available in published articles: i.e., we wanted to generate a complete picture of data gathered, of accessibility, of ease of conducting follow-ups, (for example, to ascertain whether addresses were available on the cohort), and to assess the investigator's attitude toward collaborative work.

As we publicized our survey, we received many reprints, letters, bibliographies and other written material. Our next step was to catalogue and prepare for computer processing critical aspects of the entire set of written materials. A coding scheme was devised in order to facilitate the identification of relevant projects for criminological researchers with special interests. After all the coding was done, we had the capability of listing all projects which, for example, studied biological variables; they could further be broken down by age of initial assessment, size of cohort, gender, etc. If it were of interest we could identify a list of projects of at least 500 seven-year-olds assessed for reading ability. We could also determine whether the state where the subjects reside is a state where the identification of criminological data can readily be made.

The following section presents some of the characteristics of the 380 projects in our computer file. The project descriptions from which our codes were abstracted did not always contain information on all of our coding categories. Thus the tables which follow do not always sum to 380 projects.

Characteristics of the projects.

A total of 380 projects were identified. Of the 380 projects, the great majority (279) were studies which did not involve control or comparison groups (Table 2). Moreover, most of the projects had a duration of less than 10 years with a median duration of 5.4 years. Only 24 projects had been in existence longer than 20 years (See Table 3.) A large majority of the projects (245 or 63%) studied individuals 25 years of age and under with a median age of 12.9 years at inception.

(See Table 4 for a complete distribution.) At the time of the last follow-up
the median age of the subjects had risen to 36.5 years (Table 5).

Some of the cohorts were sampled from geographical regions of the United States
(e.g., Northeast). Others drew their samples from individual states and the
District of Columbia, while 14% of the projects sampled individuals from the entire
United States. A number of the U.S. investigators studied populations in other
countries (the exact number of projects for each location is detailed in Table 6).
The type of cohort sutdied varied considerably. High school-aged and elementary-
aged children when combined accounted for 28% of cohorts (13% being elementary-
aged children). Community cohorts, pregnant women cohorts, and preschool groups
were relatively rarely studied (See Table 7). Adult (not further specified) co-
horts represent the largest single category (26%). The greatest number of project
directors (55%) selected normal subjects in contrast to a psychiatric population
(15% of data files), delinquents (4% of files), or drug or alcohol abusers (5%
of files). Table 8 includes the distribution of subject selection criteria for
the projects. Most projects (73%) studied both males and females (Table 9) with
21% looking at males only as contrasted to 6% studying females only. Almost
one third of the projects had relatively small samples (less than 100) while pro-
jects of over 500 encompassed just over one third of the total projects surveyed
(Table 10).

As noted earlier, most of the projects began with normal subjects. However,
of those projects studying deviance as an <u>outcome</u> (N=279), the largest number
(40%) focussed on psychiatric behavior; deviance in social situations was studied
by a much smaller number (17%). Drugs or alcohol use (16%) and antisocial behavior
(14%) were other important areas of concern. See Table 11.

In order to determine the variety of measures included in each data file,
tabulations were completed of the types of antecedent measures and the types of
outcome measures studied. Antecedent measures included the information collected
on the sample at the initial contact, while outcome measures were those assessed
at follow-up. Tables 12 and 13 display the frequency with which both types of
measures were used. Thus, 85% of projects assessed demographic information on
the samples as an antecedent variable while 65% included psychological measures
of personality. Of the projects, 48% examined socioeconomic/family variables.
Neurological variables were the least frequently studied (only 17%). When out-
come measures were studied, they most frequently were demographic (54%) and psycho-
logical/personality (50%). Neurological outcomes were rarely studied (8%).

Table 14 presents information on the antecedents measured in the projects
for each type of outcome observed. The table could be useful for investigators
seeking projects with specific outcome and antecedent measures. Thus if one were
interested in school outcomes for children with sociocultural or demographic
deprivation or with neurological deficits, the existence of projects with these
measures could be determined from Table 14. By correspondence with the volume
editors, the projects of interest could be identified. Tables 15-19 present
type of antecedent measure studied as a function of other characteristics of the
project such as type, age, and size of the cohort, types of subject selection
and type of research design.

The data which we have presented above in the tables were intended to be
descriptive of critical aspects of the longitudinal projects we identified. The
tables are presented in order to inform individauls interested in doing secondary
analyses regarding the wealth of suitable data currently available. As we dis-
cussed earlier, one of the intents of this volume is to encourage better utilization
of existing longitudinal files. In the chapters which follow we present detailed

description of many of these projects.

Table 1

Descriptions of Longitudinal Research

A. Correlative Longitudinal Research

   1. Normal, Representative Populations

      a. Birth Cohorts

      b. School-aged and College-aged Cohorts

      c. Military Cohorts

      d. Adult Cohorts

      e. Community Cohorts

   2. Non-representative Populations

      a. Normal Cohorts

         i. Birth Cohorts

        ii. Pre-school Cohorts

       iii. School-aged and College-aged Cohorts

        iv. Adult Cohorts

         v. Aged Cohorts

        vi. Marriage and Family Cohorts

      b. Specialized Cohorts

         i. Perinatal Damage

        ii. Mental Patients

       iii. Delinquents and Criminals

        iv. Children at Risk

         v. Disadvantaged Students

        vi. Special Children

B. Experimental-Manipulative Research

Table 2

Number of Studies with Control Groups*

	N	%
With Controls	101	26
Without Controls	279	73

*For many of the studies, control groups for specific analyses are drawn from within the cohort.

Table 3

Distribution of Duration of Studies

Duration (in Years)	N	%
0	75	22
1-10	195	58
11-20	44	13
21-30	13	4
31-40	8	2
41-50	-	-
51-60	1	4
61-70	-	-
Over 70	2	-

Table 4

Distribution of Age of Subjects at Inception of Study

Age of Subjects (In Years)	N	%
0	43	11
1-5	43	11
6-10	27	7
11-15	45	12
16-20	67	17
21-25	20	5
26-30	5	1
36-40	3	1
41-45	3	1
46-50	2	1
51-55	3	1
56-60	3	1
61-65	9	2

Table 5

Distribution of Age of Subjects at Completion of Study

Age of Subjects (In Years)	N	%
0	24	10
1-5	27	11
6-10	25	10
11-15	24	10
16-20	54	22
21-25	12	5
26-30	9	4
31-35	6	2
36-40	5	2
41-45	4	2
46-50	3	1
51-55	3	1
56-60	13	5
61-65	15	6
66-70	4	2
71-75	2	1
76-80	5	2
81-85	1	4
Over 85	13	5

Table 6

Location of Studies by State or Region

State	N	State	N
Arizona	2	New York	32
California	22	Ohio	4
Colorado	1	Oklahoma	2
Connecticut	16	Pennsylvania	12
District of Columbia	1	Rhode Island	1
Florida	5	Tennessee	1
Georgia	1	Texas	8
Hawaii	1	Virginia	3
Iowa	1	Vermont	2
Illinois	15	Washington	1
Indiana	1	Wisconsin	4
Iowa	2	West Virginia	2
Kansas	20		
Kentucky	1	Region	
Louisiana	3	Midwest	5
Massachusetts	15	Northeast	7
Maryland	13	Northwest	2
Michigan	11	Southeast	3
Minnesota	7	Southwest	4
Missouri	16	United States (Overall)	52
Nebraska	4		
North Carolina	3	Foreign	31
New Jersey	5		
New Mexico	3		

Table 7

Distribution of Type of Cohort

Type of Cohort	N	% *
Birth cohort	27	7
Pre-school	19	5
Elementary school	52	13
High School	59	15
College	38	10
Pregnant women	15	4
Military service or veterans	32	8
Adult cohort	101	26
Community/block	8	2
Other	67	17

* Percent does not total 100 as some projects fall into more than one category.

Table 8

Distribution of Subject Selection Criteria

Criterion for Subject Selection	N	% *
Normal	212	55
Psychiatric	57	15
Children at risk (Psychological)	32	8
Children at risk (SES)	20	5
Perinatal damage	10	3
Adult at risk (Sociocultural, Psychological)	7	2
Adult at risk (Medical, Physiological)	25	7
Antisocial (Delinquent)	15	4
Antisocial (Drug & alcohol abuse)	18	5
Twin	4	1
Other	25	7

* Percents may not total 100 as more than one classification may apply.

Table 9

Sex of Subjects

	N	%
Males only	76	21
Females only	23	6
Both males and females	269	73
Total number of studies including male subjects	345	94
Total number of studies including female subjects	292	79

Note: In 12 of the projects, sex was unspecified.

Table 11

Distribution of Substantive Areas of Research

Area of Research	N	% *
Educational/occupational achievement	10	3
Social/relationship adjustment	47	
Antisocial behavior	39	
Drugs/alcohol	44	
Psychiatric	106	
Neurological	11	
Health (Physiological)	19	
Other	3	1

* Percent does not total 100, and some projects fall into more than one category.

Table 10

Distribution of Sizes of Samples

Size of Sample	N	%
Under 100	119	32
101-500	123	33
501-1000	46	13
Over 1000	80	22

Table 12

Distribution of Antecedent Measures Studied

Antecedent Measures	N	% *
Physiological, Neurological	47	12
Psychological, Cognitive	148	38
Psychological, Personality	250	65
Socio/Family	184	48
Demographic	327	85
School/Educational	108	28
Medical	148	38
Deviant Behavior	148	38
Other	110	28

*Percent does not total 100 as some projects fell into more than one category.

Table 13

Distribution of Outcome Measures Studied

Outcome Measures	N	%
Physiological/Neurological	30	8
Psychological/Cognitive	104	27
Psychological/Personality	192	80
Socio/Family	116	30
Demographic	209	54*
School/Education	69	20
Medical	103	27
Deviant Behavior	119	31
Other	52	14

* Percent does not total 100 as some projects fall into more than one category.

Table 14

Distribution of Antecedents Measured for Each Type of Outcome Observed

(In Percents)

Outcomes:

Antecedents	Physiological (Neurological)	Psychological (Cognitive)	Psychological (Personality)	Sociocultural	Demographic	School	Medical	Deviant Behavior	Other
Physiological (Neurological)	18	5	4	3	3	1	6	3	1
Psychological (Cognitive)	11	21	12	11	12	14	9	8	8
Psychological (Personality)	16	18	23	20	18	16	14	18	14
Sociocultural	9	13	15	22	14	14	11	13	11
Demographic	20	22	23	23	27	22	24	24	24
School	3	9	7	8	8	22	4	5	8
Medical	18	11	10	10	11	6	25	11	8
Deviant Behavior	3	1	3	2	2	1	3	18	1
Other	1	3	3	3	5	4	4	-	26

Note: Column percents total to 100% (except for rounding error). Thus, of the projects which studied School outcomes, 16% examined psychological (personality) antecedents and 14% studied sociocultural antecedents.

Table 15

Distribution of Antecedents Measured for Each Type of Cohort Studied

(In Percents)

Type of Cohort:

Antecedents	Pregnant Women	Birth Cohort	Pre-school	Elementary School	High School	College	Military Service or Veterans	Adult Cohort	Community/Block
Physiological (Neurological)	8	8	5	3	2	1	8	3	2
Psychological (Cognitive)	7	13	18	16	13	16	2	3	8
Psychological (Personality)	19	14	15	17	21	26	11	20	23
Sociocultural	15	16	15	15	16	10	2	20	14
Demographic	17	21	22	23	25	22	33	24	25
School	3	3	9	14	13	15	-	5	5
Medical	21	18	11	8	5	4	31	10	12
Deviant	1	1	1	2	2	-	5	7	9
Other	8	5	4	3	3	5	8	7	3

Note: Column percents total to 100% (except for rounding error). Thus, of the
projects which studied high school cohorts, 16% examined sociocultural
antecedents and 5% studied medical antecedents.

Table 16
Distribution of Antecedents Measured for Each Subject Age Group (At Inception of Study)
(in percents)

ANTECEDENTS	Age at Inception												
	0	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-50	51-55	56 & Over
Physiological (Neurological)	7	4	1	2	3	3	8	-	-	9	-	10	6
Psychological (Cognitive)	11	16	15	10	10	4	-	-	-	9	17	10	13
Psychological (Personality)	13	15	18	19	18	18	17	-	-	9	-	20	15
Sociocultural	16	12	13	13	14	13	-	-	-	9	-	10	13
Demographic	19	21	20	22	25	28	42	-	-	27	33	20	23
School	4	10	11	11	8	1	-	-	-	-	-	-	2
Medical	15	10	7	5	11	7	42	-	-	27	50	20	11
Deviant	7	7	11	13	12	10	8	-	-	9	-	-	13
Other	7	4	5	5	11	15	-	-	-	-	-	10	4

Note: Column percents total to 100% (except for rounding error). Thus, of the projects which studied 6-10 year olds (at the inception of the study), 11% studied school antecedents and 7% studied medical antecedents.

Table 17

Distribution of Antecedents Measured for Each Size Category

(in percent) -

SIZE:

ANTECEDENTS	Under 100	101-500	501-1000	Over 1000
Physiological (Neurological)	4	3	2	4
Psychological (Cognitive)	12	10	6	10
Psychological (Personality)	18	18	14	15
Sociocultural	11	14	13	11
Demographic	20	21	25	25
School	7	7	8	8
Medical	11	8	11	12
Deviant	11	11	9	8
Other	6	8	12	7

Note: Column percents total t. 100% (except for rounding error). Thus, of the projects of under 100 subjects, 20% studied demographic antecedents and 12% studied ps. chological (cognitive) antecedents.

Table 18
Distribution of Antecedents Measured for Each Subject Selection Criterion
(in percent)

SUBJECT SELECTION CRITERIA:

ANTECEDENTS	Normal	Psychiatric	Children at risk (Psychological)	Children at risk (SES)	Perinatal damage	Adult at risk (Sociocultural, Psychological)	Adult at risk (Medical, Physiological)	Antisocial (Delinquent)	Antisocial (Drug & Alcohol abuse)	Twin	Other
Physiological (Neurological)	2	4	7	2	11	5	4	–	4	10	5
Psychological (Cognitive)	12	9	12	14	15	5	6	–	9	–	14
Psychological (Personality)	18	25	19	16	11	24	10	–	23	20	17
Sociocultural	14	16	17	15	15	18	8	–	20	–	12
Demographic	26	22	20	23	17	24	32	50	23	50	21
School	10	6	6	14	4	–	1	–	1	–	9
Medical	9	8	13	8	19	21	35	–	14	20	14
Deviant	1	8	4	3	6	3	–	50	5	–	4
Other	8	1	1	3	2	–	4	–	–	–	5

Note: Column percents total to 100% (except for rounding error). Thus, of the projects which selected perinatal damage subjects, 11% studied physiological (neurological) antecedents and 15% studied psychological (cognitive) antecedents.

Table 19

Distribution of Antecedents Measured for Each Type of Design
(in percent)

TYPE OF DESIGN:

ANTECEDENTS	Epidemiological	Follow-up	Prospective	Retrospective	One-time	Longitudinal or cross-sectional	Methodological/theoretical	Other
Physiological (Neurological)	7	3	4	2	4	4	1	4
Psychological (Cognitive)	8	12	12	8	9	10	10	16
Psychological (Personality)	10	19	18	21	22	18	23	17
Sociocultural	12	14	15	13	14	13	14	13
Demographic	27	25	24	26	24	27	30	20
School	8	8	9	11	7	9	7	7
Medical	24	11	11	11	10	12	7	10
Deviant	5	3	3	2	5	2	2	6
Other	–	5	5	6	4	7	5	6

Note: Column percents total to 100% (except for rounding error). Thus, of those projects which were retrospective, 13% studied sociocultural antecedents and 26% studied demographic antecedents.

References

Douglas, J.M.B. Early hospital admissions and later disturbances of behavior and learning. Developmental Medicine and Child Neurology, 1975, 17, 456-480.

Silverman, J., Berg, P.S., & Kantor, R. Some perceptual correlates of institutionalization. Journal of Nervous and Mental Disease, 1966, 141, 651-657.

Wadsworth, M.E.J. Delinquency in a National Sample of Children. British Journal of Criminology, 1975, 15, 167-174.

Wadsworth, M.E.J. Delinquency, pulse rates and early emotional deprivation. British Journal of Criminology, 1976, 16, 245-256.

Wolfgang, M.E., Figlio, R., & Sellin, T. Delinquency in a Birth Cohort. Chicago, Illinois: University of Chicago, 1972.

Chapter 2

A Brief History of North American

Longitudinal Research

Sarnoff A. Mednick, Ph.D., Dr. Med.

Birgitte Mednick, Ph.D.

University of Southern California

Social Science Research Institute

A Brief Outline of the History of
Longitudinal Research

Sarnoff A. Mednick, Ph.D., Dr. Med.
Birgitte Mednick, Ph.D.

University of Southern California
Social Science Research Institute

Before the 1920's, longitudinal research was represented by scientifically curious individuals (typically parents) who measured and described the growth and development of their children. The earliest such report mentioned in the literature is the study made by Montebeillard of his son's growth between 1759-1777 (Buffon, 1799). Several writers have described or referred to such early investigations (Scamnion, 1927; Baltes and Nesselroade, 1980; Wall and Williams, 1970).

Some 19th century methodological work centering on growth and development contained quite sophisticated consideration of methodological problems in longitudinal research. For example, Quetelet's (1842) research design discussions could be useful reading for modern researchers.

While these careful chronicles of individual development can be cited as spiritual forebears, what we now understand as longitudinal research had its origins in the second decade of this century. Sontag (1971) reports that two men, Beardsley Ruml and Lawrence K. Frank played a pivotal role in stimulating research on the development of human life. Under the aegis of the National Research Council, a series of conferences on life history research in child development was held which helped launch the classic studies initiated in the 20's. The conferences inspired an engagement in understanding the biological and social forces which interact to form "the whole child."

The first of the classic, longitudinal projects was initiated by Terman in 1921. His aim was to study the physical, mental and personality development of intellectually superior children. (For a current report on this work, see chapter 53). This project was distinguished from the projects which immediately followed in that it was focussed on gifted children. The other projects have a more general orientation seeking to understand the development and growth of the normal child. Following the Terman Gifted Child Study, projects developed at the Merrill Palmer School in Detroit and at the University of Colorado Medical School (both in 1923) and at the University of Minnesota (1925). In 1928, the Berkeley Growth Study began; in 1929, the Berkeley Guidance Study (see chapter 41), the Fels Research Institute Project and the Harvard Longitudinal study were launched. The final project in this series was the Oakland Growth Study initiated in 1932. While these projects had different emphases depending on the investigators who initiated them, they all shared a common interest in the explication of the physical, mental and personality development of the child. They were consciously multi-disciplinary making detailed anthropometric measures, studying skeletal development by X-ray, assessing personality with the TAT and Rorschach and recording details of family circumstances and interactions. The samples ranged in size from 61 to 309 subjects and were (with the exception of the Berkeley Guidance Project)

almost always highly selected, white, middle-class volunteers. The investigators felt more secure about continued cooperation with such subjects.

How might one characterize these pioneering studies? Lester W. Sontag, one of the founders of the Fels Research Institute Project, characterized them by the following five points:

1. They were somewhat global in their approach.
2. They had elected a methodology without delineating many specific problems.
3. They lacked stated hypotheses.
4. They were all committed to research efforts which would for years, be heavily devoted to data collection.
5. They were committed to making a variety of kinds of observations and measurements for which there were not yet immediate research questions of designs. (page 989)

It is clear that such projects would never be given a high priority score today by our government study sections. It is, however, also clear that had the original longitudinal researchers begun their studies more recently they would themselves have chosen to alter the methodology of the studies in several respects. It is an inherent problem in all longitudinal research that the knowledge, technology and preferred research paradigms of a field will be subject to dramatic changes during the duration of a project. Thus, with the benefit of hindsight, longitudinal researchers will always find themselves wishing they had done things differently in the initial project phases. If one at this point in time should make an overall evaluation of the adequacy of the methodology used in the pioneer longitudinal studies, one would have to say that given the state of the art in developmental psychology at a time when the studies were begun, the original investigators were very skillful in their choice of procedures. For example, the global, rather than problem-specific, focus and the lack of initial hypotheses, which led to extensive collections of many diverse categories of data, appears to have increased the long-term potential of the studies. Due to diversity of the data banks, researchers of very different backgrounds and theoretical orientations have been able to find data relevant to their particular interests in the study files. Examples of the diversity of the problems analyzed by use of the data from these projects are presented by the works of Kagan and Moss (1962), McCall, et al. (1973), and Block (1980). Clearly, the fifth point made by Sontag, concerning the initial collection of data categories not related to specifically formulated research questions and methodologies, further enhanced the research application of the original longitudinal projects; for example, this increased the chances of longitudinal data being available for analyses of research questions which became of interest to psychologists after the completion of the initial data collection phases of these projects.

A continuing criticism of these projects in the early years was their lack of productivity. However, this must be considered a highly unjustified criticism born out of lack of understanding of the fundamentally different goals and objectives which were guiding the efforts of the pioneer researchers in this area. It was the opinion of these researchers that developmental psychology had gone as far as it could using the more traditional and less-time-consuming research paradigms. They contended that further understanding of affective

and cognitive development of the etiology of behavior problems and of the extent to which developmental prediction is possible on the bases of early assessments could best be determined by longitudinal, prospective research (MacFarlane, 1963). Thus clearly with this set, high levels of publication activity during the early data collection phases of the studies was neither considered important nor very desirable. That is, before the types of analyses for which the longitudinal investigations were planned, could be undertaken, data had to be collected over a rather lengthy period of time and the system for organizing, keeping up and protecting the data files for the future had to be developed. The originators of the projects assigned these tasks, rather than publication, the highest priority during the early years. Their willingness to make this kind of extensive investment of time and effort deserves lasting gratitude and respect from their contemporaries as well as future colleagues. Due to their effort psychologists have now access to invaluable collected data banks, the research potential of which is still far from exhausted.

At this time, the concern regarding the productivity of longitudinal research projects has been effectively countered by the important and frequently quoted research reports which have emerged from these studies since the mid-fifties. In addition to authorships by the original longitudinal investigators, this literature presents many instances of younger researchers who, in analyses of a current problem, were fortunate enough to be able to draw the benefits of the groundwork done during the many preceding years of data gathering.

An important factor which facilitated the exploitation of the rich data banks of the longitudinal studies was the establishment of the United States Public Health Service as a granting agency. This agency helped guide scientists to consider more problem-oriented studies and more focussed aims that could be more clearly stated. Support was provided to researchers who showed intentions and abilities to carry out such research with the data banks of the child development institutes begun in the twenties.

The projects of the twenties did not spawn a rush of similar projects in the United States or Europe for some years. It was not until 1948 that the next project along the lines of the original 1920's project was initiated by Sybil Escalona at the Menninger Foundation in Topeka, Kansas. This study focussed on the influence of family life and personal trauma and on the dynamics of personality development. This rather long time span during which no new longitudinal studies were begun (i.e., after the first group from the twenties) may very likely be ascribed to the many practical difficulties encountered by such projects and by the slow payoff in terms of results and publications. Perhaps the great depression of the 30's also had a damaging effect.

The Current Status of Longitudinal Research

In the modern era, projects have become much more focussed and/or much larger. A good example of a relatively small project with focussed interests is the New York University School of Medicine study of behavioral development (Thomas, et al., 1960; Thomas, et al., 1963). Thomas and coworkers were specifically interested in the question of the relationship of primary (constitutionally in the relationship of primary (consti-

tutional) reactivity in infancy and the long-term meaning of this type of
reactivity to psychological growth. The focussed nature of this study has
meant that a series of important questions were always available to help in
structuring data analyses and publications (see Chapter 38). It also meant
that the precious initial assessment techniques were problem-focussed
and were therefore appropriate for the questions posed later in the project.

Project TALENT, (see Chapter 24) the Educational Testing Service Project
and the U.S. Collaborative Project are other examples of major efforts in the
United States which were originated to answer relatively specific questions.
Some of these have evolved into multipurpose projects as the subjects aged and
began to evidence a variety of outcomes (e.g., delinquency, mental illness,
etc.....).

The larger longitudinal studies of the more modern era are exemplified
by the giant British birth cohort ushered in by the 1946, 1958, and 1970 studies.
These studies chose as their population all individuals born in a specific
week in Great Britain. All perinatal data available on the included subjects
were collected and related to later follow-up studies of the total cohort or
of selected subsamples (Davis, Butler and Goldstein, 1972; Douglas, 1964;
Chamberlain and Chamberlain, 1975). This type of design obviously results in
a study sample which is representative of the total population from which it
is drawn. However, for research purposes other than census-type analyses, the
non-uniform conditions under which the initial data collection took place in
these studies constitutes a methodological disadvantage. The perinatal
data were collected and recorded by varied sources of hospital personnel,
midwives attending clinics or home deliveries or non-medical witnesses
to a delivery. Thus more missing data and less standardized data collection
seems to be the price paid for the representativeness of the English cohort
and similar samples. Clearly, the mentioning of the problems associated
with the data collection procedures in these studies should not in any way
be interpreted as a down-grading of the merits of the type of study in general.
The obvious value associated with the increased generalizability of findings
resulting from the study of representative samples speaks against that
interpretation. Rather the problems were pointed out in an attempt to illustrate
populations involving intensive and standardized data collection, and of the
larger studies of representative populations with resulting diminished control
over the experimental procedures. Both types of studies represent unique
capabilities for answering research questions. The studies of representative
populations are useful in describing population trends and incidence rates of
outcomes in a population without control of interacting variables. The studies of
specially selected samples are important in estimating relationships in specific
groups homogenized in terms of social or medical conditions. Depending on the
nature of the research questions one or the other or a combination of both
may provide the most adequate paradigm.

Perhaps the newest breed has been the high-risk projects. Such projects are
highly problem-oriented. They study individuals at high-risk for later
deviance (for example, children of schizophrenics); such subjects and controls
are assessed relatively early in their lives before signs of the target deviance
are manifested. After some years when some proportion of the sample does

become mentally ill or criminal one can then analyse the data of the
initial intensive assessment and subsequent follow-ups and determine
how the deviant individuals distinguished themselves from their more
fortunate co-subjects at the time of the initial assessment. The
oldest of these projects is the Copenhagen High-Risk Project which has
been following a population of children of schizophrenics for the past
20 years (Mednick, Schulsinger and Griffith, 1981). Another interesting
type of study assesses a biological measure early in childhood and
observes eventual outcomes. This may be illustrated by the Mauritius
Project which assessed the autonomic nervous system functioning of
a total population of 3-year-old children in Mauritius (Mednick, Venebles,
Schulsinger and Cudeck, in press).

This volume, together with its companion volume on European longi-
tudinal research (Mednick and Baert, 1981), present a fair picture of
the current state of this field in the developed western nations. At
this point in time, the research support climate does not encourage new
longitudinal research initiatives. Perhaps it is time for careful
retrenchment and more dedication to secondary analysis of existing studies.

Baltes, P.B. and Nesselroade, J.R. in J.R. Nesselroade and P.B. Baltes (eds.) Longitudinal research in the study of behavior and development. New York: Academic Press, 1979.

Block, J.H. and Block, J. The role of ego-control and ego-resiliency in the organization of behavior. In W.A. Collins (Ed.) Minnesota Symposia on Child Psychology. Vol. 13. New York: Lawrence Earlbaum, 1980.

Buffon, (1799) Cited in J.R. Nesselroade and P.B. Baltes (eds.) Longitudinal research in the study of behavior and development. New York: Academic Press, 1979.

Chamberlain, R., Chamberlain, G., Howlett, B. and Claireaux, A. British births, 1970, London: Heinemann, 1975.

Davis, R., Butler, N. and Goldstein, H. From birth to seven, London: Longman, 1972.

Douglas, J.W.B. The home and the school. London: McGibbon and Kee, 1964.

Kagan, J., and Moss, H.A. Birth to maturity. New York: Wiley, 1962.

Macfarlane, J.W. Perspectives on personality and change from the Guidance study. Vita hum., 1963, 7, 115-126.

McCall, R.B., Hogarty, P.S. and Hurlburt, N. Transitions in Infant Sensorimotor Development and the Prediction of Childhood IQ. American Psychologist, 1973, 27, 278-88.

Mednick, S.A., Schulsinger, F., and Griffith, J. Children of Schizophrenic mothers: The Danish high-risk study. In Schulsinger, F. and Mednick, S.A. (Eds.) Longitudinal Research: Methods and uses in behavioral science, Boston: Martinus Nijhoff, 1981.

Mednick, S.A., Venables, P.H., Schulsinger, F., and Cudeck, R. The Mauritius project: An experiment in primary prevention. In Goldstein, M.J. (Ed.) Preventive intervention in schizophrenia: Are we ready?, Government Printing Office, Washington, D.C., in press.

Quetelet, A. A treatise on man and the development of his faculties. Edinburgh: William and Robert Chambers, 1842.

Scammon, R.E. The first seriation study of human growth. American Journal of Physical Anthropology, 1927, 10, 329-336.

Sontag, L.W. The history of longitudinal research: implications for the future. Child Development, 1971, 42, 987-1002.

Thomas, A., Chess, S., Birch, H.G., Hertzin, M.E. and Korn, S. Behavioral individuality in early childhood. New York: New York University Press, 1963.

Chapter 3

Overview of Longitudinal Research:

Methodological and Practical Problems

Michele Harway

Sarnoff A. Mednick

Birgitte R. Mednick

University of Southern California

Most of the chapters which follow present individual researchers' experiences with the longitudinal approach and descriptions of long-term longitudinal data files. In some instances, investigators have also highlighted difficulties encountered in the completion of the projects. While it is not our intention to gloss over the problems which face individuals who invest time and energy into the collection of longitudinal data, we do have some suggested strategies and solutions. Moreover, it continues to be our contention as well as that of other proponents of the longitudinal approach that there are some problems which can be studied effectively only by the longitudinal method; the study of development and issues having to do with long-term consequences of events such as the incidence and development of disease or deviant behavior are judged to be most appropriately researched with the longitudinal approach.

In the sections which follow, we discuss methodological and practical problems facing longitudinal researchers.

Practical Difficulties

Cost.

The most frequently cited objection to longitudinal research has been the cost associated with the data collection over an extended period of time and with the long term support of a research staff. Let us examine, however, the assumption that longitudinal research is costly. While the difinitive data on this question are not currently available -- to our knowledge no one has actually costed out a longitudinal study to support or refute the allegations -- there is some evidence that this research method is in fact less expensive than others. Inflationary factors aside, the biggest cost in a longitudinal project is the initial data collection, with the tasks of sampling, variable conceptualization and refinement, overall design of the research, and initial instrument development to be completed prior to the initial assessment. Moreover, staff training needs to be done most thoroughly at the beginning of the project and apparatus needs to be built or purchased. In addition, the initial data collection is usually completed on an entire cohort, while later follow-ups may limit themselves willingly or not to subsamples. Thus costs for the data collection alone are usually higher at the beginning of the study, as are overall costs. If only one additional follow-up assessment is planned ten to twenty years later, then pro-rated costs per year are likely to be quite low. When intermediate assessments are made, the costs are likely to rise somewhat.

Yet, consider what the costs would be if some other research approach were utilized. In order to obtain the same amount of data as collected in a longitudinal project using some other type of research, it would most likely be necessary to study several separate samples, each examining some variable or set of variables at one point in time. With each separate study, new subjects would need to be selected and assessed, new apparatus would be needed and the research staff would need to be trained for each new data collection procedure. Thus the costs for other types of research almost always exceed those for equivalent longitudinal studies.

When the costs of longitudinal research are compared to the costs of cross-sectional studies, (in cases where the two procedures are adequate for the problem to be researched) the major consideration has to do with the size of the sample required. The sampling error in longitudinal studies is related only to the representativeness of a single sample whereas in cross-sectional studies each additional sample adds to the sampling error, with the result that a longitudinal sample may be kept smaller than the combined size of a number of cross-sectional samples, leading to a cost savings.

With a new sample needed to be selected for each cross-sectional study and with the high cost of sample selection, it is evident that this procedure would lead to higher expenditure for the same pay-off than would the longitudinal approach. In addition to costs of sample selection, the cross-sectional approach would also require background data to be collected again on each sample and staff to be trained each time.

A second cost consideration is related to the fact that much longitudinal research is multidisciplinary and multi-purpose. That is, such studies often include a wide number of variables relating to social and medical deviances and covering a variety of fields such as psychology, medicine, sociology, etc. Because of the breadth of coverage of these studies it is usually possible to conduct many analyses and substudies with diverse foci. The individual cost for each sub-area to be investigated is substantially reduced as compared to collecting the same amount of information with other research methodologies. Such multi-disciplinary longitudinal projects may also lend themselves to secondary data analyses to explain outcomes (such as mental illness, social deviance, etc.) which were not the original focus of the project. It seems obvious that undertaking such secondary analyses would result in considerable savings.

Publication Record.

A second criticism of longitudinal research is directed at the publication record of researchers using this method. Some of the early longitudinal investigators were more interested in studying the fate of their subjects than in publishing findings of the study. The belief remains today that there is a paucity of published material on the longitudinal approach. The fact is that the literature abounds with reports on longitudinal research. This is less obvious than it might be because many researchers prefer to publish papers in their own discipline since status and prestige are enhanced more by such a procedure than by publishing in either an interdisciplinary journal or one in a different discipline altogether. Thus, the politics of professional recognition are such that they may create one of the greatest barriers to the multidisciplinary interest of longitudinal research. A corollary of this is that researchers may be unaware of longitudinal research in closely related disciplines other than their own, since they are less likely to search for literature in other disciplinary journals.

Thus, locating longitudinal research projects is no easy job, since no one publication or field is likely to reference all of the ongoing or recently completed work. Another chapter in this volume describes the process by which the editors of this volume located the projects described herein. The disciplinary nature of longitudinal research suggests that a great deal more longitudinal research

is published than is generally acknowledged and that an interdisciplinary volume such as this may begin to bridge the gap and to inform researchers in a wide variety of fields.

Funding.

Another practical difficulty mentioned at some time or other by all longitudinal researchers is the difficulty in obtaining and/or maintaining funding. As a consequence of the belief that longitudinal research is more costly than other research approaches, funding agencies have been reluctant to grant continuing support to longitudinal research. Moreover, the compartmentalization of government-granting offices into discipline-specific funding groups has made it difficult for multi-disciplinary projects to get support. Government funding agencies in particular are subject to political pressures which may constrain their ability to fund long-term research projects, even those with potentially important practical implications. Thus, many longitudinal researchers find that government agencies are not supportive of their planned projects. Other researchers, who have received government funding for years, may find their support abruptly cut off at year nine of a twenty year project (for example) because the agency feels that the project has gone on "long enough". All of these factors do present discouraging picture for the future of funding for longitudinal research. Counterbalancing this outlook are several encouraging signs: There have been indications from within government that there is an interest in establishing study sections focussing on longitudinal research within some of the larger funding agencies. A further encouraging trend is that these same agencies have been increasingly seeking to fund projects which involve secondary analyses of existing data files rather than projects which depend on new data collection. Thus, it may be possible to sustain funding from these agencies by developing broad data files which lend themselves to multidisciplinary research and secondary analyses or by identifying existing files with the goal of doing secondary analyses. Thus, another intent of this book -- to identify existing data files which may be appropriate for further analysis on different topics from those specified by the original investigator -- may result in prospects of collaboration where funding is not as much of a problem.

Timeliness and inclusiveness.

The collection of longitudinal data has also presented practical problems due to the fact that the passage of time leads to changes in instrumentation or theory. Thus, information needed to pursue a hypothesis developed at a later time may not have been collected at the inception of the project because its possible significance was not perceived. For example, testing hypotheses about Legionnaire's dis-

ease from a long-standing medical data file may prove difficult because relevant information may not have been collected. This is likely to have occured because the disease was not identified when the medical file was developed. Thus, changing definitions and improvements in laboratory conditions and technology have made researchers aware of conditions and illnesses not previously identified. In many cases, the exploration of those conditions depends upon earlier data which were not collected. Or changing conditions may affect the criteria under consideration in such a way that data at two points in time become noncomparable. For instance, to ask an individual about his or her political orientation (i.e. radical, moderate, or conservative) during the 60's may not yield information comparable to that gathered on the same topic in the 80's, since concepts about radicalism and ocnservatism have changed substantially in the last two decades.

The problem of changes in hypotheses and interest areas is a real one. The longer the study continues the more such change is likely to occur. Whether the value of what has already been collected diminishes is dependent on the nature of the data. Thus, good planning involves designing a longitudinal study which anticipates the possibility of spotting unpredictable influences and is flexible enough to allow for the possibility of change and the exploration of the impact of unanticipated variables. In addition, it is important to keep the theory which guides the research fairly general, such that it may change over time. It is also a good planning technique to include a variety of measures in the data collection instruments and to record these in raw data form so that they might be reanalysed at a later date and interpreted in the light of empirical and technological advances. Finally, a well-planned longitudinal project might include, as Gruenberg and LeResche (1980) have suggested, an advisory board of distinguished scientists from many different disciplines. Such an advisory board, especially one independent of the political constraints of the funding agency, could provide the expertise necessary to develop an inclusive and flexible data base and later could suggest the analysis of issues from a variety of perspectives.

Data Storage.

Longitudinal researchers are often concerned with the difficulties inherent in the storage of their data. The sheer volume of information collected and the clerical time involved in processing the data, can indeed become a problem. More than in any other type of research approach, however, accuracy and precision of data collection and storage are essential in longitudinal studies. The advent of the computer age has made the storage of data a simpler task. Yet, because it is not usually possible to regather or reconstruct data that have been lost or improperly processed, it is critical that the investigator institute procedures that allow for periodic evaluation and quality control of the data collection phase of the project.

Staffing.

The final area that has presented a practical problem in doing long-term longitudinal research, has been the maintenance of a research staff. Few individuals, except for perhaps the principal investigator, are likely to remain on staff for the duration of the project, (sometimes 20 or 30 years). Since training people is a financial burden, shifts in personnel have often been costly to longitudinal projects. On the other hand, periodic changes in personnel may not be unhealthy. In fact, the occasional addition of new intellectual blood may very well lead to the inclusion in the data base of more variables from a variety of disciplines. This is likely to happen more often with shifts in personnel than if personnel were to remain stable. In addition, new personnel may reanalyze the existing data in new and creative ways not considered by the ongoing staff members.

Methodological Issues

Selection of appropriate variables for the outcomes of interest.

A methodological problem closely related to the practical problems of timeliness and inclusiveness of the data is that having to do with the selection of appropriate variables to study for the outcomes of interest. Age of subjects at the time of data collection and at planned follow-up ages is an important factor to consider in the choice of variables to be included in a longitudinal research project. The literature in developmental psychology has demonstrated that the degree to which successful prediction of short and/or long-term outcomes of a given variable may be obtained is highly influenced by variations in subject age. In a prospective study initiated with subjects who have reached adolescence, a rather high degree of predictability can be expected between the adolescent measures of personality traits and cognitive ability and similar measures obtained later in adulthood. Measures obtained on children during elementary school ages show considerably less correlation with later measures (Thomas, Chess and Birch, 1968; McCall, Appelbaum and Hogarty, 1973). The lowest levels of successful long-term prediction by far have been obtained in prospective longitudinal studies begun during infancy and early preschool years (Sameroff, 1979).

In this discussion, while the examples which follow are selected from studies beginning in infancy, generalizations may be made to research starting with older subjects. Our discussion of this issue centers around three types of infant variables: perinatal and neonatal trauma, social variables and infant assessment measures.

The risk types of variables: perinatal and neonatal physical traumas have been shown to have a significant, though generally not strong effect on developmental outcomes during early childhood. (Sameroff, 1979).

In addition, some evidence exists for a sleeper effect being operative in the long-term outcomes associated with this category of infant predictor variables. (Mednick 1978; Pasamanick & Knobloch 1966). The research evidence seems to suggest that information regarding perinatal and neonatal events will have the most predictive value if the planned follow-up ages are during preschool years; follow-up during elementary school years is likely to show no significant main effects of perinatal variables.

The level of predictability of later outcomes on the basis of perinatal and neonatal constitutional data tends to be greatly improved by addition of data describing conditions of the developmental environment. Thus, obtaining such information should be assigned high priority in studies aimed at predicting towards long-term outcomes as a function of perinatal and neonatal events.

The research literature examining the effects of a second kind of variable, that of static environmental or social variables (Sameroff, 1979; Bradley, Caldwell and Elardo, Bronfenbrenner, 1974) suggests that this category of predictors has no measurable impact on development and behavior during the first year of life. Sometime during the second year, most often around 18 months of age, the negative effects of less optimal environmental conditions begin to appear, and from then on, the negative effect becomes progressively more and more pervasive (Kagan, Lapidus, and Moore, 1978). Thus, it seems important that in addition to static measures of the social environment, measures describing the characteristics of the caretaker and of the caretaker-child interaction be included in prospective longitudinal research on infant samples. Since inclusion of these categories of predictor variables has been shown to improve the predictability of early and later emotional and cognitive outcomes to a highly significant degree.

Inclusion of variables describing caretaker characteristics and interaction style must be seen as especially crucial in studies beginning in early infancy. At this point of development, observations of caretaker characteristics and infant interaction patterns present the possibility for determining, to some degree, the role played by each of the actors in forming the long-term interpersonal relationship. As first pointed out by Bell (1968), the pattern of interaction developing between a child and its caretaker is by no means developed through a unidirectional process; that is, through the infant's responding to the caretaker's behavior. Rather, the pattern of interaction developing in an infant-caretaker dyad depends to an equal degree on the characteristics and typical reactions of each member of the dyad.

By observing an infant-caretaker dyad from earliest infancy, some insight into the etiology of the child-adult interaction observed at later ages may be obtained. In contrast, the etiological information to be obtained about the interactional pattern itself as well as its impact on concurrent child characteristics must be seen as limited if one studies the interaction of adolescents with their parents. As an example, watching parents of a juvenile delinquent

interact with the youngster and observing an unloving parental style will not allow us to conclude that this parental characteristic might have a causal relationship to the delinquency. It is also possible that such a child has caused the parents an extraordinary amount of difficulty or trouble (perhaps starting in infancy) by possessing characteristics that made caring for and raising the child a very frustrating task. In such circumstances, the observed parental style could be interpreted more correctly as an effect, rather than as a cause. Obviously, interaction patterns observable between a child and his or her parents may never actually be divided in such a simple fashion into causes and effects because they have all been formed through the continued interaction of parental and child characteristics.

The results from the literature on prediction of shorter- and/or longer-term outcomes on the basis of infant assessments, the third set of variables show such prediction to present a rather disappointing low level of success, particularly in the areas of cognitive functioning and personality traits (Thomas, Chess and Birch, 1968; St. Claire, 1978; Rosenblith, 1964, 1973; Sameroff, Kratchuk and Bakow, 1978; Kagan, Lapidus and Moore, 1978; Horowitz, Sullivan and Linn, 1978; Crano, 1977; Corah et al, 1965; Ucko, 1965; McCall, Hogarty and Hurlburt, 1973; Sameroff and Chandler, 1975.

The lack of predictability of later functioning from infant measures should not be interpreted to mean that infant assessment data are inappropriate for use in longitudinal research. It is, however, important to note that confining the selection of measures to those considered to be infant measures of the traits or areas of competence that are to be measured at later subject ages will yield disappointing results. The recommendation to be extracted from the research seems rather to be that as wide a range of measures as possible, including varied measures on the infant, the environment, and their interaction, shows greater promise for successful prediction.

It should be pointed out that the increasing level of predictability of measures obtained at later ages in childhood is very likely to be caused by the fact that such measures possess higher qualitative continuity with the final outcome measures studied. In young adulthood the measured characteristics cease to change in quality from test to test, thus resulting in a high stability of measures obtained from late adolescence onward.

As mentioned earlier, some of the conclusions made about infant research seem generalizable to research begun with older subjects. The following points seem among the more important to mention in this

category:

1. The planned follow-up ages of the subjects have implications for the kinds of variables to be included and the level of prediction to be expected in longitudinal research.

2. The importance of biological measures in prediction is heightened if such measures are analyzed in interaction with environmental measures.

3. The predictive power of environmental measures as well as the possibility of gaining insight into the mechanisms through which environmental effects on development are mediated is dramatically improved by inclusion of nonstatic environmental data. Examples are data based on direct observation describing the characteristics of significant persons in the environment and of the patterns of interaction developed between the subject and these persons.

Sampling:

In addition to the variable selection and design considerations discussed above, a frequently cited methodological difficulty in longitudinal research has to do with sampling. The sampling of longitudinal research has been critiqued from several perspectives: Respondents or probands for longitudinal studies often have been chosen because of their cooperativeness and availability. This of course makes the sample not representative and subject to selection biases. However, it certainly is possible to obtain random samples or to structure the study around birth or community cohorts. In all instances, cooperativeness is required unless unobstrusive methods are utilized.

Another type of sampling problem reported is caused by changing population patterns. The population may thus have changed so much because of immigration or emigration that it no longer bears much resemblance to the original population assessed. Researchers assessing this type of cohort must plan their study on the population as it is projected to be at the time of analysis rather than the population as it is at the inception of the study. Moreover, judicious follow-up of the original sample may remedy the problem to a substantial extent.

Follow-up:

Still another kind of problem has to do with follow-up difficulty and the consequent effects on the representativeness of the resulting sample. A frequent occurence is that over long periods of time, attrition of the sample occurs and the comparison of the initial and final groups is difficult to specify. However, proper planning can also prevent this problem from plaguing the project. Some researchers (Wall & Williams, 1970; Crider, Willits & Bealer, 1972) are reporting

successful follow-up procedures of 80-90%, the last figure for a study of over 20 years' duration. Efficient administrative follow-up can also reduce gaps in the data owing to the absence of subjects at particular data collection sessions. It is also possible to weight the sample for nonresponse bias since antecedent characteristics are available on all subjects.

It must be emphasized that difficulties of sampling and gaps in data are not confined to longitudinal work. Other techniques are subject to the same problems, for example, house to house surveys may suffer in representativeness due to the absence from the home of persons sampled.

Repeated measures:

One type of problem, specific to longitudinal research, has to do with the experimental effect of testing. In testing and retesting the same individuals, the initial testing always carries the possibility of affecting respondents' behavior in such a way that they would behave differently on retesting than if there had been no original testing. For example, if a researcher were interested in the effect of point of view in a film on students' attitudes, a research study might be devised in which students were presented with a film presenting one side of the argument in an extreme and yet convincing manner. To test the effect of the film, the researcher might administer a pretest and a post-test of students' attitudes. If there is a shift in the students' attitudes between pretest and post-test, can the researcher legitimately attribute this change to the film? Would such a change still occur if the subjects had not been sensitized to the film by virtue of having been pretested? This is one problem induced by repeated testing. Another problem is commonly known as the "practice effect" when an increase of scores occurs on the post-test by virtue of the fact that the subjects have practiced on the pretests. Experimental designs do exist which correct for the experimental effects of testing (Campbell & Stanley, 1966). With longitudinal research however, the problem is likely to be magnified as testing or measurement may consist not just of testing and retesting, but potentially of an infinite number of data collection waves. The continued and repeated testing of the same individuals carries the risk of modifying their behavior in unknown ways. The problem is not unsolvable as longitudinal designs have been developed which in effect provide control conditions (Campbell & Stanley also describe such procedures, as do many other researchers, including Mednick, in press).

There is a related point which may be made regarding repeated testing: Continuous contacts increase the rapport between researchers and subjects, with the result that it may be easier to collect certain types of sensitive data.

Data collection intervals:

On the other end of the continuum, lack of frequency of data collection may also provide a problem. For example, characteristics of the subjects may change at a rate very different from that provided for in the data collection schedule. Thus, when trying to relate outcomes (such as criminality, for example) to environmental or personal changes (such as changes in family composition) which have occured during two data collection points, it may be less possible that usual to infer which of several changes is causally related to the outcome in question. Again, good planning in the selection of a measurement interval would deal with this potential problem. In selecting a measurement interval, a researcher may want to consider:

a. the rate of change of the characteristics in question

b. the degree to which the sample is linked to some cohort defining event

c. the frequency of occurence of events between measurement

d. the extent to which intermediate events occur simultaneously for all subjects

Scale construction:

Another issue may stem from the fact that research in the social sciences frequently depends on questionnaires containing dichotomous items. Such dichotomous items often are translated into dichotomous variables. These are generally not appropriate for analysis with multivariate techniques which are useful for sequential analysis or complex types of data analysis (e.g. analysis of covariance structures, path analysis). Investigators can plan to develop a body of data, making possible the construction of scales which combine the dichotomous items and produce scores with appropriate range and variance. Useful in this type of scale construction is Cronbach's alpha technique.

Confidentiality:

The final area which needs to be mentioned here is that of confidentiality and the related issue of data linkage. While the first issue is common to all research, it presents a particular problem with longitudinal research because of its long-term nature. Thus, the confidentiality and protection of participants must be maintained for longer periods of time with resultant complexities. Moreover, where large amounts of data have been collected on large numbers of respondents, there is always the danger that unscruplous individuals will violate the rights of participants or that the data will be subpoenaed thorough the courts. Thus, it is imperative that participants be protected and some organizations such as the American Council on Education have gone to great extremes to do so. (See the chapter on ethics

and confidentiality for more details).

Data linkage can also present a problem particularly in cases where separate files have been set up for names of respondents and data collected. Since follow-up must either be precoded with subject number or coded after they have been returned to the research organization, the file linking names and subject numbers must be obtained thus temporarily increasing the risk of exposure of individual subjects. Selection of highly ethical and responsible personnel in this area is essential to maintaining the confidentiality of respondents.

This overview of methodological considerations indicates that while engaging in longitudinal research complicates somewhat the research task, none of these methodological problems is insurmountable. Much of the criticism levelled against longitudinal studies may have been warranted in the past but proper planning may avoid many of the methodological defects we have touched on in this chapter.

References

Bell, R.Q., 1968. A Reinterpretation of the Direction of Effects in Studies of Socialization. Psychological Review. 75, 81-95.

Bradley, H.R., Caldwell, B.M. and Elardo, R., 1979. Home Environment and Cognitive Development in the First 2 Years: A Cross-Lagged Panel Analysis. Developmental Psychology. 15(3), 246-250.

Bronfenbrenner, U., 1974. A Report on Longitudinal Evaluations of Preschool Programs. DHEW Publication No: (OHD) 75-25.

Campbell, D.T. & Stanley, J.C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally and Company, 1966.

Corah, N., Anthony, E.J., Painter, P., Stern, J.A. and Thurston, D.L., 1965. Effects of Perinatal Anoxia after Seven Years. Psychological Monographs. 79(3) No. 596.

Crider, D.M., Willits, F.K. & Bealer, R.C. Locating People in Longitudinal Studies: A Research Report and Suggested Guidelines. Bulletin 77, University Park, Pa. : Pennsylvania State University, 1972.

Crano, W.D., 1977. What Do Infant Mental Tests Test? A Cross-Lagged Panel Analysis of Selected Data from the Berkeley Growth Study. Child Development. 48, 144-151.

Horowitz, F.D., Sullivan, J.W. and Linn, P., 1978. Stability and Instability in the Newborn Infant: The Quest for Elusive Threads. In A. Sameroff, (ed.), Organization and Stability of Newborn Behavior: A Commentary on the Brazelton Neonatal Behavior Assessment Scale. Monograph of the Society for Research in Child Development. 43 (5-6).

Kagan, J., Lapidus, D.R. and Moore, M., 1978. Infant Antecedents of Cognitive Functioning: A Longitudinal Study. Child Development. 49, 1005-23.

McCall, R.B., Appelbaum, M.I. and Hogarty, P.S., 1973. Developmental Changes in Mental Performance. Monograph, Society for Research in Child Development. Vol. 38 (3).

McCall, R.B., Hogarty, P.S. and Hurlburt, N., 1973. Transitions in Infant Sensorimotor Development and the Prediction of Childhood IQ. American Psychologist. 27, 278-48.

Mednick, B., 1978. Intellectual and Behavioral Functioning of 10-11 Year Old Children Who Showed Certain Transient Neurological Symptoms in the Neonatal Period. In S. Chess and A. Thomas (eds.) Annual Progress in Child Psychiatry and Child Development. New York: Brunner/Mazel.

Mednick, S.A.   Methods of Prospective, Longitudinal Research. In S.A. Mednick & A.E. Bart (eds.).   An Empirical Basis for Primery Prevention: Prospective Longitudinal Research in Europe. Oxford University Press, 1979.

Pasamanick, B. and Knobloch, H., 1966.   Retrospective Studies on the Epidemiology of Reproductive Casualty: Old and New.   Merrill-Palmer Quarterly.   12, 7-26.

Rosenblith, J., 1964.   Prognostic Value of Behavioral assessments of Neonates.   Biologia Neonatorum.   6, 76-103.

Rosenblith, J., 1973.   Relations Between Neonatal Behaviors and Those at Eight Months.   Developmental Psychology.   10, 779-92.

St. Clair, K. , 1978.   Neonatal Assessment Procedures: A Historical Review.   Child Development.   49, 280-92.

Sameroff, A., 1979.   The Etiology of Cognitive Competence: A Systems Perspective.   In R.B. Kearsley and I.E. Sigel (eds.), Infants at Risk.   New York: Wiley.

Sameroff, A. and Chandler, M., 1974.   Reproductive Risk and the Continuum of Caretaking Casualty.   Review of Child Development Research.   187-244.

Sameroff, A., Krafchuk, E.E. and Bakow, H.A., 1978.   Issues in Grouping Items from the Neonatal Behavioral Assessment Scale. In A. Sameroff (ed.), Organization and Stability of Newborn Behavior: A Commentary on the Brazelton Neonatal Behavior Assessment Scale. Monograph of the Society for Research in Child Development. Vol. 43, (5-6).

Thomas, A., Chess, S. and Birch, H., 1968.   Temperament and Behavior Disorders in Children.   New York, New York University Press.

Ucko, L.E., 1965.   A Comparative Study of Asphyxiated and Non-asphyxiated Boys from Birth to Five Years.   Developmental Medicine and Child Neurology.   7(6), 643-57.

Wall, W.D. & Williams, H.L.   Longitudinal Studies and the Social Sciences.   London: Heinemann, 1970.

Chapter 4

Longitudinal Methods in the Study of Development*

L.N. Robins

Longitudinal Methods in the Study of Development*

L.N. Robins

What Longitudinal Studies of Children Have Taught Us

The quetsions that longitudinal studies have successfully answered are those dealing with 1) the prevalence of various types of children's problems in the population at large, in children of various ages, in girls as compared with boys, and in children in different social classes; 2) the correlations between various childhood problems; 3) the likelihood that particular problems or personality characteristics will persist or remit; 4) the identification of treatment or intervention effective in changing the likelihood of the persistence of problems; 5) the association of family and other aspects of the childhood environment with the appearance of problems in childhood and with the continuation of childhood problems into adulthood; 6) the association of childhood problems and traits with the emergence of difficulties as adults; and finally, 7) the demonstration that many of the environmental variables and types of intervention commonly believed to influence children's outcomes do not actually do so.

The Prevalence of Problems in General Populations

The first decision to be faced in producing estimates of problem rates in a general population concerns the cutoff points defining pathology. Often we use definitions that simply decide a priori to consider a certain percentage of the population as "abnormal." To designate that certain percentage as abnormal, we first construct a scale along which a particular characteristic varies, fix the scale's midpoint at the population mean and call "pathological" those cases that fall more than two standard deviations above or below that midpoint. In other words, we create a standardized scale. The IQ test is the most common example of such a scale, where a score of 100 has been set as the

midpoint and individuals more than two standard deviations below the mean are considered retarded. Sometimes researchers forget that IQ scores mean nothing more than that, and solemnly report that half the children in their sample have low (below 100) IQs! It would be surprising indeed if they did not! National reading evaluation tests, like the IQ, are standardized scales designed so that a fixed proportion will be defined as backward readers. Once so defined, the lowest group is offered remedial efforts. It is important to remember that the determination of how many poor readers will qualify for help is entirely an a priori one. It merely reflects how much investment the society plans to make in improving reading; it is not an absolute measure of reading disability! Thus the finding that 6.5 percent of 10-year-olds on the Isle of Wight (Study T1) are backward in reading is chiefly useful as a baseline against which children from specially disadvantaged or advantaged backgrounds, with or without psychiatric problems, can be contrasted.

Estimates for psychiatric disturbances do not have such a predetermined distribution of scores in the general population, in part because no standardized measuring devices like IQ and reading tests exist. Indeed, estimates of impairment rates vary widely in normal populations, depending on the criteria applied. For example, Rutter (Study T1) found a prevalence of 8 percent for psychiatrically disturbed 14-year-olds when he based his estimate on teacher and record data, but after he obtained questionnaire responses, his estimate rose to 21 percent, the increase due principally to self reports of "often feeling miserable." In the National Child Development Study (Davie, Study P), first grade teachers were asked to identify children who would hardly ever sit still. They named 11 percent of the boys and 5 percent of the girls. Teachers were also asked to assess children on the Bristol Social-Adjustment Guide, and scored 14 percent in the "maladjusted" range. In their sample of 136 children followed from birth to adolescence, Thomas and Chess (Study V) identified a third with serious enough emotional problems at some time in their childhoods that help was sought. The Study V sample was especially predisposed to seek psychological counseling, being New Yorkers largely from the professional class who had been induced to participate in the study in part by the offer of professional help to the child. Nonetheless, the authors judged the problems presented not to be trivial.

These widely diverse estimates make it clear that different methods yield different results. Yet, however one looks at the problem, there seems to be a considerable number of children thought by themselves or others to have psychological problems. It also seems clear from Rutter's followup (Study T1) that problems increase between the ages of 10 and 14. Using the same criteria at both ages, Rutter identified 11 percent as

psychiatrically disturbed at age 10, and 16 percent at age 14. The addition was chiefly among girls, and chiefly in the area of neurotic disturbances, rather than in learning or conduct problems.

Boys not only show problems earlier, but they have a different spectrum of problems. Among those found to predominate in boys at age seven are hyperactivity, reading problems, enuresis, hostility, withdrawal and restlessness. Young girls exceed young boys only in anxiety (Davie, Study P).

In addition to the increase in psychological disturbance with adolescence, there are changes in attitudes that may be upsetting to family and school. As children reach high school age, they typically become less conventional in their attitudes, less convinced that school achievement is to be sought, more independent, less attached to their parents and less religious -- or so a recent study that followed them for four years found (Study J). These findings confirm lay views of adolescence as a period of exploring new values and separating oneself from parental views. It should be remembered, however, that this study represents only a brief historical period, and one in which there was a marked upsurge of rebellion in adolescents. To demonstrate that these changes are typical of adolescence in general, replications in different historical periods would be necessary.

Correlations Between Problems

One of the most important contributions of followup studies has been the demonstration of the association between problems of different types. Finding associations between behaviors we had previously perceived as discrete suggests that they may actually be part of a common syndrome, or if not, may share some of the same causes. When behaviors are associated, we begin to look at them in a broader context rather than as single symptoms. Thus for some underlying process rather than as single symptoms. Thus for many years, there was a great deal of speculation in the literature about the causes of alcoholism. Theories coming from psychoanalysis focused on drinking as an _oral_ activity, and looked for its causes in infantile oral experiences. Newer research, however, has found that early drinking occurs in the very same children who are delinquent, drug users and sexually advanced (Studies J, R2). If children indulging in forbidden sexual contacts are the very same children indulging in forbidden drinking, it makes it difficult to argue that drinking represents fixation at a pregenital stage of sexual development.

All the behaviors we have just identified as occurring in the same children -- drinking, sex and delinquency -- all appear to come

from a single conceptual realm: they all represent resistance to authority. There are also examples of related behaviors from other single conceptual realms. For instance, Rutter (Study T1) has found low IQ and poor reading to be associated with poor performance on school achievement tests, associations that we could have anticipated, since school performance should be a function of both native ability and successful learning. The more interesting results, however, are associations discovered between problems that seem to come from different realms. Poor reading is associated not only with IQ and attendance, but also with general maladjustment (Studies P and Q) and psychiatric disorder (Study T1), particularly conduct disorders. Poor reading is additionally associated with neurological disorders such as epilepsy and cerebral palsy. One connecting link may be IQ, which is also associated simultaneously with conduct disorder and neurological abnormality (as well as with bronchitis, speech impairment and left-right confusion) (Study T1). However, IQ is not the sole explanation since children with conduct disorders read even more poorly than their IQs would lead one to expect (Study T1).

Another reliable association found through longitudinal studies is that between enuresis and conduct disorders (Studies R1 and T1) -- again an association that is by no means obvious. The association is an intriguing one, since enuresis can be alternatively interpreted as a delay in maturation or as a refusal to conform to family norms.

Not only problems but desirable behaviors have been shown by longitudinal studies to be strongly correlated. Douglas (Study Q) found, not surprisingly, that well adjusted, nonneurotic students did best in the "eleven-plus," the examinations that English schools used to assign students into academic or general education tracks. Gifted children (Study U) were found to score above average on health, physical attractiveness and athletic ability. While these findings are consistent with later results, they were surprising at the time when there was believed to be a law of compensation at work that made bright children athletically or socially inept. The only suggestion that there may be some slight penalty to be paid for having unusual assets is Kagan and Freeman's (Study E) observation that for girls only high IQ is associated with aggression, independence and less dating behavior in high school. (This study was also done a number of years ago. Given the changes since then in views of proper roles for women, it is no longer so clear that these are adverse consequences.)

Persistence and Termination of Problems

Discovering the prevalence of problems in general populations and the correlations between different problems is a byproduct of longitudinal

studies, but similar results could have been obtained in cross-sectional research. The special role of longitudinal studies is in observing change over time. Observations over time lend themselves to two broad goals: discovering which childhood problems persist and which remit, and discovering predictors of future problems.

We noted previously that the prevalence of problems increases between ages 10 and 14, and that the ratio of boys to girls decreases as girls begin to catch up with the boys' higher early rate. This increase in overall rates and equalizing of the sexes could be due either to the disappearance of young boys' problems and the onset of new types of problems common to both sexes in adolescence, or to the persistence of young boys' problems and the addition of girls' problems in adolescence.

Rutter (Study T1) found that the latter alternative was the one that actually occurred. Very few of the boys diagnosed as having a conduct disorder at 10 had recovered from it by 14, nor did those with reading disability overcome it. Many studies (Studies C, G, R1 and R2) agree that delinquents have a high risk of continuing their criminal behavior, although the frequency of rearrest begins to taper off in the middle 20s. These studies also agree that continuity of crime is greatest when the juvenile crime was more serious, when the first delinquency occurred at a very young age and when the juvenile offense led to institutionalization. Kagan and Moss (Study E) found that aggressive boys remained aggressive men, and indeed agressivity was the most stable personality trait in their sample. Langner (Study M), following New York City children for five years, found fighting, conflict with parents, delinquency and "mentation problems " (i.e., academic problems) all to continue or worsen over the five-year period. Heber (Study I), following the offspring of mentally deficient mothers as control subjects for his treatment project, found that their slightly low IQs in early childhood dropped increasingly with age. Thus there is ample evidence that the conduct and learning problems typical of the disorders of early childhood are often persistent, although it is also clear that many children with early problems improve.

Langner (Study M) presents the best evidence about the ages at which these various handicaps stabilize enough to allow them to be used predictively. He found that antisocial behavior becomes stable at about age 10, while conflict with parents and mentation problems are stable much earlier -- about age six. On the positive side, early high IQ and early school achievement are also stable, as Terman (Study U) found for IQ and Kagan (Study E) found for achievement. The Berkeley Growth Study (Study A) reported trait stability beginning about the age of school entry. Children showing destructiveness, attention seeking, shyness, somberness and jealousy at ages six and seven were likely to show the same traits in adolescence.

While findings are remarkably consistent among studies about the stability of antisocial, aggressive and intellectual problems after the first school years, studies differ in their findings about the stability of neurotic traits. Douglas (Study Q) found that children least well adjusted, i.e., "neurotic," tend to deteriorate in school performance after age eight. Rutter (Study T1) found some consistency for neurotic traits, but much less than for conduct problems. Of Isle of Wight children judged neurotic at ten, half were healthy at 14, as compared with only one-fourth of the conduct disorders. Still, neurotic children were more likely to be affected at 14 than were children free of symptoms at ten, particularly if they had been seriously enough affected to require hospitalization. Roff (Study S) found little relationship between neurotic symptoms as recorded in a child guidance clinic and success in military service, suggesting that the neurotic symptoms either remitted fully or became mild. Langner's (Study M) findings may be the key to bringing some order out of the chaos of contradictions about childhood neurotic symptoms. He found that over a five-year interval, a number of "neurotic" symptoms did get better: dependence, repetitive motor behavior, delusions and hallucinations and social isolation. However, he points out that only one of these, social isolation, had been associated with the overall impairment score in his initial survey. He found that anxiety, the classic "neurotic" symptom, did not become a stable characteristic until adolescence. His results would indicate that _most_ neurotic symptoms in children are both mild and transitory. The more serious and incapacitating ones, and those arising in adolescence, may be a good deal more stable.

Studies of general populations of children often group together all symptomatic children whose problems are neither in the aggressive, disciplinary problem realm nor learning problems into the residual category "neurotic." In this "catchall," along with children with anxiety, fears and overdependence, there is a very small group with serious mental illness. Unlike the rest of the children in the "neurotic" category, these seriously disturbed children have extremely poor prognoses. Childhood psychoses, as shown in followup studies by Rutter and Lockyer (Study T2), Bender (1973), Annell (1963), Eaton and Menolascino (1967) and Eisenberg (1957) are very serious disorders indeed, rarely remitting and extremely disabling.

Although studies differ in their categorizations of childhood problems, there is great consistency in the finding that personality, achievement patterns, social skills and aggressiveness are largely set well before adolescence into patterns that are likely to be liabilities or assets for a lifetime. Thus the treatment of children with serious impairment in these areas is important not only to provide immediate relif but for the child's long term adjustment.

The Identification of Successful Treatments

Treatment evaluation is the form in which the followup study can function as a true experiment. Ideally, children are randomly assigned to treatment and control groups, the treatment process is carefully specified and evaluation of outcome is blind and unbiased. While this ideal design is familiar and well accepted, attempts to carry it out have been surprisingly rare and those studies that have made the attempt often find that dropouts badly mar the equality of treatment and control groups that was initially achieved by random assignment. In addition, the outcome criteria used have often been so general that one can only tell in the positive studies that some positive effect has occurred, not its nature. Where outcome measures have been explicitly defined, it is often found that improvements have been achieved in psychological test scores or teachers' impressions, while measures such as academic achievement and rates of arrest for delinquency, which would be more meaningful for the child's future, remain distressingly unaffected (Gittelman-Klein & Klein, 1976).

In all, there have been few studies showing significant long term effects of treatment (Robins, 1973). Even worse, there have been some showing significant adverse effects. McCord (Study C) reports results of a treatment program in which counselors tried to help boys and their families in every way possible, including tutoring, medical attention, summer camps and recreational programs. Followed 30 years later, the treatment group had more early deaths, more high blood pressure and heart trouble, worse jobs, more job dissatisfaction and more repeated arrests. Another study with disturbing results (Study O) followed any three years later children who had been assigned at age four to four types of preschools and to a control group. The experimental children in all programs were advanced compared to controls on entering first grade, but by grade two, treated children all showed declining IQ scores and children from three of the four programs had lower scores than did the untreated controls. Girls particularly seemed to suffer detrimental effects. Thus, it appears that psychological as well as physical treatment can result in unanticipated noxious effects.

One well designed study by Heber (Study I) presents more cheering results. The offspring of 20 women with markedly low IQs were placed in day care before six months of age, and intervention was continued until age six. In day care, the children were offered an unusually intense learning environment. At age seven, the index children continued to show remarkably high IQ levels when compared both with a control group of offspring of women with equally low IQs and with their own older siblings.

This last study, if its results are confirmed in larger samples, will join a small select group of proven successful treatments for children's psychological problems. Followup studies of behavior modification techniques have produced impressive short term results in home and classroom, and desensitization has effectively reduced childhood phobias (Achenbach, 1974). It is interesting that treatment has shown its most enduring positive effects with regard to those childhood symptoms -- fears -- that have a high rate of spontaneous remission.

The Association of the Family and Other
Childhood Environments with Problems

The family factors most often proposed as predictors in followup studies include family size, broken homes, illegitimacy, adoption or foster placement, socioeconomic status, supervision by parents, attitudes of parents toward the child, parental expectations for his achievement, behavior problems in the parents and siblings and psychiatric disorder in the parents. The difficulty in assessing the role of these family variables is that they are all strongly intercorrelated. Families of low socioeconomic status have more illegitimacy, more breaks, more very small or very large sibships, offer less supervision (in part because poor mothers have to work), have less expectation that the child will receive higher education, include more parents who are criminal, mentally retarded, alcoholic or schizophrenic. Because of these intercorrelations, a longitudinal study that chooses one or another of these characteristics to show that it predicts bad outcomes in the child could often come to the same conclusion had it picked any other from this list of variables. It is only when some of the descriptors are held constant while the relationship of another to children's problems is examined that we begin to get a hint as to what the causal mechanisms might be and therefore what aspects of family pathology are most important.

Clearly there is a marked association between families with some or all of these characteristics and both school failure and conduct disorder in offspring. For instance, Farrington and West (Study X) have found an association between low social status and delinquency. Low status has also been implicated by Davie (Study P) in poor reading, poor mathematics skills and poor social adjustment as measured by withdrawal, dependency, hostility and restlessness. Rutter (Study T1) also found an association between social status and adjustment, which disappeared when he controlled for IQ. However, Douglas (Study Q) found the association of low social status with school problems not to depend entirely on the inheritance of low IQ, since the association continued even when

the child's IQ was kept constant. The association might have additionally depended on the poorer school attendance of his lower-status children. Further evidence that the family's social status has an effect independent of IQ comes from the study of gifted children (Study U), which found that social status in childhood continued to predict adult occupational status even in this group where all IQs were very high.

Status in childhood continues its effects into adulthood. Havighurst (Study H) found that lower class status of school children predicted poor adjustment in their early 20s. Roff (Study S) found that men of lower class backgrounds were more likely to be diagnosed as having conduct disorder in the military service, but were not more likely to be discharged for neurosis.

Coming from very large families has been shown to depress IQ scores, although by only a small amount, with the effect most pronounced in the later-born children (Record, et al., 1969) and in children closely spaced (Zajonc, 1976). The negative effects of large families on other aspects of adjustment have been reported repeatedly. Davie (Study P) found large sibships associated with poor reading and poor adjustment in first grade; Douglas (Study Q) also found them associated with poor school achievement. Similarly Rutter (Study T1) reported a relationship of large families with deviance at age 10, as did Farrington and West (Study X) with delinquency. Robins (Study R1) found coming from large families added to the risk of sociopathy in adulthood.

Next to social class and family size, the family variable most often reported associated with children's problems is the broken home. The association of the broken home with official delinquency is a well established one, recognized in the early part of this century and reaffirmed more recently by the Gluecks (Study G) and by Farrington and West (Study X). Broken homes are associated as well with behaviors of an antisocial type that do not come to official attention, as Rutter (Study T1) has shown. Roff (Study S) has similarly shown an association between broken homes and bad conduct discharges from the military. Among gifted children (Study U), broken homes were related to low occupational status and to a second generation of broken marriages.

Families can be broken in various ways and for various reasons; the child may be sent away to relatives, foster parents or an adoptive home; or a parent may leave, die or be institutionalized. A child may be born to a mother and father who never lived together. Following the break, the child may live in a variety of settings: with a parent, with relatives, with foster parents, or in an institution. Many studies have now shown that the type of broken home, as reflected in its cause and consequences, is an important predictive factor. Thus Robins (Study

R1) found that the effect of the broken home in predicting sociopathy was entirely attributable to the problems in the parents that led to their divorce, separation and institutionalization. Quarreling antisocial parents who remained together were no improvement as far as their offspring's adult diagnosis was concerned.

Douglas (Study Q) has shown the importance of the type and time of family breakup. Early breaks (before the child is six) caused by divorce or separations were found strongly associated with sons' delinquency and with daughters' having an illegitimate child. The National Child Development Study (Study P) compared illegitimate children who remained with the mother and those adopted. Most of those adopted ended in middle class homes, while most of those remaining with their mothers lived in lower status one parent homes. Further, adoptive homes offered a high level of parental interest in school achievement, while maintaining equally small sibships as did the setting with the natural mother. Apparently as a result of these relative disadvantages, children kept by the mother had lower school achievement and more adjustment problems than did the adopted children. Despite the assets of the adoptive home as compared with the average "intact" family in terms of income and small sibships, adopted boys, but not girls, nonetheless had more problems with coordination, restlessness and adjustment than did children from intact families, although they did better in verbal achievement.

The fact that adopted children do as well as legitimate children living with their natural parents on most measures of ability and adjustment suggests that adoption may be the most successful preventive measure available. Since adopted children share with other illegitimate children many of the high risk indicators -- young mothers, poor prenatal care and small birth weight -- their normality at followup is probably attributable to their better environments, although screening does prevent the adoption of those illegitimate children with the worst prognoses.

The illegitimate child left with his mother is clearly at a major disadvantage, no doubt in part because of objective deprivation, but apparently also in part because he was unwanted. The dismal outcomes for children born to mothers who do not want them was underscored by a followup (Study F) of children born to mothers who had been refused abortions. As compared with the next child in the birth registers, they had more delinquency, more psychiatric care and were more often educationally subnormal.

A final interesting but controversial discussion about parents' effects has grown out of the attempts to study the genetics of schizophrenia by comparing the offspring of schizophrenics with the offspring

of control mothers. Not only has an excess of schizophrenia been identified in these children, but also an excess of alcoholism and drug abuse (Schulsinger, Study N). To what extent this represents the genetic contribution of the schizophrenic parent and to what extent it reflects the effect of assortative mating between schizophrenic women and antisocial men or the effects of family breakup due to the mother's illness is still unsettled.

Siblings as well as parents can have important effects on outcome. Farrington and West (Study X) and Robins (Study R2) have found that having delinquent siblings increases the risk of delinquency. Terman and Oden (Study U) found that bright and successful siblings increased the probability of achieving a high occupational status for bright children.

In addition to family influences, longitudinal studies have explored cultural influences on children's outcomes. Rutter (Study R1) found rates of disorder twice as high in an inner London community as on the Isle of Wight, and Wolfgang (Study Y) found a strong impact of race on delinquency rates in Philadelphia.

There is a difficulty in assessing the impact of cultural influences because ethnic status and residence are highly correlated with family variables. Black boys in Philadelphia (Study Y) and inner city Londoners (Study T1) are much more likely to have broken homes, arrested parents, low socioeconomic status and large sibships than are the groups with whom they are contrasted. As mentioned previously, when populations vary profoundly along such dimensions, matching and statistical controls can never fully overcome the differences between them. Therefore, even attempts to show the effect of one variable holding another constant may not be adequate to prove a relationship, though they are often adequate to disprove one. Thus, Robins (Study R1) was able to show that low social class in childhood was no longer an important predictor of adult psychiatric status once the parent's level of antisocial behavior was taken into account.

Among the family and background factors that have been found to remain predictive of children's outcomes when factors correlated with both background and outcome were held constant are: large families predict poor school achievement, holding social status constant (Study Q); low social status predicts poor reading and social adjustment holding constant family size, mother's age and child's legitimacy (Study P); low social status predicts delinquency holding constant family size (Study X); criminality in biological parents of adoptees predicts their criminality, holding constant the criminality of the adoptive parents (Study D1).

How children's infractions of the law are handled by the legal system seems to play an important role in the child's later adjustment. While no particular form of treatment of delinquency has generally been found to reduce the likelihood of recidivism, studies of the diversion system (i.e., whether infractions lead to appearance in court and incarceration) have generally shown that more formal and severe handling of juvenile offenses has been associated with worse outcomes (Studies R1 and 2, S, Y). It is always difficult in a nonexperimental study to be certain that delinquents treated with greater severity were not initially more seriously delinquent or from more disadvantaged homes. But efforts to hold seriousness of the offense and background factors constant have not been able to wipe out an apparent independent effect of the way the child was handled. Study Y did not find this effect for the first offense, but did for subsequent offenses.

The Association of Childhood Problems and Traits
With the Emergence of Later Difficulties

Some of the characteristics portending later childhood difficulties can be detected at birth. Children born both too early and too late have been shown by the National Child Development Study (Study P) to have poorer school achievement and more physical incoordination at age seven. These results are well substantiated in a number of studies, including that by Barker and Edwards (1967) where premature infants were also found to be at high risk for a number of physical disabilities -- epilepsy, cerebral palsy -- which in turn are associated with later adjustment problems.

Hospital admissions in early childhood were found by Douglas (Study Q) to forecast a variety of problems in adolescence: conduct disorder, reading problems and delinquency. In general, however, traits prior to school age except in the severest form such as childhood psychosis and extremely low IQ have not been shown to predict later outcome (Study A). Chess and Thomas (Study V) attempted to link temperament in the first year with the appearance of psychiatric problems in later childhood. The only significant relationship they found was with activity levels and this finding itself was unstable since the same trait in the second year of life was no longer significantly related to later psychiatric problems.

Once children reach school age, however, stability greatly increases. A variety of indicators of trouble in elementary school are strongly related to both adolescent and adult outcomes. Among substantial predictors of later delinquency are recklessness or "daring,"

resistance to teachers' authority, fighting and aggressiveness, unpopularity with peers, low IQ, school failure, excessive absence, hyperactivity, reading retardation and poor grades (Studies G. H. R2, X and Y). Similar behavior traits are associated with later alcoholism (Studies C, R1 and 2, U), poor work adjustment (Studies B, H, R1 and 2) and adult smoking (Stewart & Livson, Study A).

Perhaps the most general of all childhood predictors is low IQ. It has been shown to influence not only delinquency, reading problems and likelihood of completing school, but also smoking (Study Z), early marriage and poor work adjustment (Study H), and later psychiatric problems (Studies L, T2 and W). Both Watt (Study W) and Lane (Study L) found some relationships between low IQ and schizophrenia. Lane in addition has demonstrated that schizophrenics' premorbid childhood IQs showed less correlation with their siblings' IQ than do normal children's, suggesting a possible neurological basis for the disorder demonstrable early in life.

Almost as broadly predictive of later difficulty as low IQ is early aggressiveness. Havighurst (Study H) found it predicted generally poor adjustment in the early 20s, as well as poor grades in high school, smoking and nervousness. Robins (Study R1) found that it predicted a host of adult problems, including criminality, low socioeconomic status, poor job performance, marital instability, drinking problems, poor performance in the military, isolation from relatives, impoverished social relationships and geographic mobility. One of its most common early effects is to increase the risk of delinquency (Study X).

Delinquency in turn has profound implications for later adjustment. The Gluecks (Study G), Roff (Study S) and Robins (Study R1) in follwoing delinquents into adulthood found not only a high rate of criminality but also poorer health, vagrancy, conduct problems in the military, financial dependency, drinking, marital friction, rejection of traditional social relationships such as church and voluntary organizations, promiscuity, gambling -- indeed increased rates of problems along almost every dimension of poor adult adjustment.

There is no question, on the basis of these longitudinal studies, that school failure, delinquency and aggressiveness are all predictors of a host of later adjustment problems. Work by Jessors (Study J) clarifies our understanding of how such different kinds of behaviors in childhood can have so similar and generalized an effect. The Jessors found that one could predict transitions of school aged children into drinking, loss of virginity, marijuana use and delinquency about equally well, whichever behavior appeared first. In addition, a drop in achievement orientation, a loss of interest in religion, a shift toward peers and away from parents as the source of values all predicted that in the

next year, whichever of these behaviors had not yet occurred would be likely to be tried. They interpreted these findings to mean that there is a normal development during adolescence in the direction of opposing adult values and that the selection of specific acts is largely a matter of chance. Robins (Study R1) also found some evidence for a generalized predisposition to deviance in that the variety of previous deviant behavior was the best predictor of further problems. However, there also seemed to be some more specific relationships, so that early poor school performance predicted school dropout and marijuana use predicted trying other drugs, even when the variety of deviant behavior was constant.

The Berkeley Growth Study (Study A) sought predictors of adjustment at various ages, the last report dealing with adults at age 30. For girls the principal predictors of adult maladjustment were finicky eating, excessive modesty and excessive dependency in the preadolescent period (age 11 to 13). For boys there were two alternative patterns leading to poor psychiatric health. One was the introverted, shy and somber boy; the other a quarrelsome, negativistic boy with temper tantrums.

At present, evidence for predictors of an antisocial adult adjustment are much more substantial than are predictors of adult neurosis. Roff (Study S), for instance, searching child guidance clinic records for predictors of a discharge from service for psychoneurosis, found only having been referred to the clinic by a physician and having poor peer relationships, while there are many predictors of rejection and discharge from service for bad conduct: police referral, low IQ, being older, fewer years of schooling and again, poor peer realtionships. While Robins (Study R1) and Rutter (Study T1) both found some increased risk for disturbance in later life associated with children's neurotic symptoms, the stability and strength of these relationships were considerably less impressive than for relationships between antisocial behavior in childhood and later problems.

The search for the later significance of childhood symptoms has helped to clarify important diagnostic issues. Rutter and Lockyer (Study T2), for instance, were able to show that childhood psychosis does not turn into adult schizophrenia, and Robins (Study R1), Watt (Study W) and the Dallas studies (Michael, et al., 1957) were able to show that, contrary to clinical impressions, male schizophrenics as children are not typically shy and retiring, but on the contrary are described by teachers and in child guidance clinic records as less conscientious, less agreeable and somewhat more aggressive than other children. Robins (Study R2) and Johnston (Study Z) were able to show that drug abuse was not simply one aspect of a general antisocial pattern, since unlike truancy from school, delinquency and incipient alcoholism, it was not related to lower class status, broken homes or poor achievement in elementary school.

Some Important Negative Findings

In the preceding pages we have emphasized the most solid associations found in longitudinal studies between family and social settings and children's behaviors, and between children's characteristics in early childhood and their later success or problems. To complete the picture we must emphasize the many reliable negative findings -- findings of no association between background and later outcome.

1. Violent and aggressive behavior patterns do not appear in adults if they have been absent in childhood -- except of course in the context of a specific physical or psychiatric disorder like mania, drug intoxication or temporal lobe epilepsy (Study R1).

2. Social and cultural environments which are not reflected in the under-the-roof culture of the child's home (e.g., in family discord, parental deviance or overcrowding) have little impact on delinquency. Thus the child living in a well functioning home located in a high delinquency area is about as unlikely to become delinquent as is a child living in a low delinquency area (Studies R2, X).

3. Family breakup per se is not an important predictor of delinquency. There is no elevation of rates when the breakup is due to death. When parents are deviant, delinquency is equally common when parents do and when they do not separate (Study X).

4. Children of working mothers do not perform worse in school, at least not if the mother's working is delayed until the child starts school (Studies P, R2).

5. Parental neurotic problems do not lead to delinquency in children (Study Q).

6. Most of the difference in delinquency rates between schools can be explained by the kinds of children who attend them, rather than by the school's own environment (Study X).

7. Pre-schizophrenics' IQs do not decline over the childhood years (Study L).

8. The contribution of the biologic parents to schizophrenia is not the result of the schizophrenic mother's providing a poor intra-uterine environment, since the risk for offspring of male schizophrenics is no different from that of female schizophrenics (Hanson, et al., 1976).

9. The association between parents' and children's deviance is not principally explained by the child's modeling his behavior on the parents' since a) children separated from the parent have rates of problems similar to those with the affected parent in the home (Study R1); b) adopted children without criminal biological parents do not have increased rates when the adoptive parent is criminal (Study D1); and c) younger delinquent siblings account for as great an increase in the risk of delinquency as do older delinquent siblings (Study X).

10. Neurotic symptoms do not protect children against becoming delinquent, although they do not increase the risk (Studies R1 and X).

11. Birth complications are not associated with delinquency or schizophrenia (Studies N and X).

12. Shy, withdrawn boys are not at high risk of schizophrenia (Studies R1 and W; Michael, et al., 1957).

Negative findings are in some ways the most important contribution that followup studies of children can make. All societies invest a large part of their resources in efforts to improve the mental health of their population. Followup studies of children can indicate areas in which it would be a mistake to expect investments to have large payoffs and can point to areas in which none of our intervention methods seems successful, so that we know where better methods must be developed. The negative findings listed above have some very practical implications about what not to expect in planning prevention and treatment. For instance, removing juvenile delinquents from the neighborhood into detension centers will probably neither improve their own prospects nor protect their peers from "infection." We know this because institutionalization is no cure for delinquency and because family problems are better predictors of delinquency than are neighborhood rates. Nor will removal of children from problem parents wipe out the inheritance of problems in one generation. We know this because children adopted away from criminal, alcoholic and schizophrenic parents early in life still have an increased risk of problems.

Positive findings of followup studies have provided clear indicators of which children are at greatest risk, so we know for whom we need to discover better forms of intervention. It is the negative finding that antisocial behavior virtually never occurs in adulthood if absent in adolescence, however, that suggests that if we do discover an effective intervention, it can be proven effective fairly promptly since there is not likely to be massive backsliding in adulthood if we can successfully interrupt the conduct disorders of childhood.

A Word for the Future

Followup studies of children have taught us that children's problems are highly interrelated among themselves and highly predictive of a wide variety of adult problems. These strong intercorrelations mean that when we do develop successful techniques for prevention and intervention, they are likely to have wide ranging beneficial effects. If we could only treat learning problems successfully, for instance, we might simultaneously get a reduction in delinquency and psychiatric disorders in children, not to mention in all the correlated adult problems such as crime, financial dependency and alcoholism.

Unfortunately the results of followup studies can rarely be translated into suggestions for intervention. While they have taught us a great deal about predicting the occurrence of problems in children, we have learned very little about predictors of the course of those problems once they occur.

Epidemiologists have provided a useful formula which accounts for the prevalence of any disorder: Prevalence = Incidence x Duration. Thus far almost all our attention has been expended in discovering predictors of incidence (i.e., new cases), not of duration. Demographic factors such as age, sex, race, status and family type have been found to be such powerful predictors of incidence, but they have not been shown to be of much help in explaining duration. Predictors of incidence are valuable in planning prevention strategies, but they give us no clues as to how best to plan treatment, which devotes itself to reducing duration. Monitoring duration in longitudinal studies is more difficult than studying incidence. We must be able to determine not only whether the disorder is present or absent but when it disappeared and under what circumstances. The potential contribution of longitudinal research in discovering determinants of the duration of childhood problems is so enormous that it should inspire the next generation of research efforts.

Notes on Studies*

A. Berkeley Growth Study

Study Type: Real-time prospective. Index Cases: Subjects were drawn from two initially distinct studies: (1) the Berkeley Guidance Study, which selected healthy, full term babies born in Berkeley over an 18-month period (N = 248) and assigned matched pairs as treatment or

---

*This bibliography was prepared with the great assistance of Kathryn Strother Ratcliff, Ph.D.

control subjects; and (2) the Oakland Growth Study, which selected fifth graders from five Oakland elementary schools (N = 212). Length of Followup Interval: For the Guidance Study sample information was recorded from birth to approximately age 30 years and for the Oakland Growth Study from fifth grade to approximately 37 years. Many interim studies used shorter intervals. Interval Measurement: Semi-annual in the Growth Study and at least yearly to age 14 in the Guidance Study, but with widely fluctuating numbers of cases contacted at different ages. Data Sources: Teacher ratings; interviews with subjects, teachers, parents, siblings, spouses; physical assessments; psychological tests; observations of child's behavior; school report cards; sociometric ratings by schoolmates. Case Recovery: Varied widely from study to study. Block and Stewart and Livson both used both samples: Block studying 171 (51%) out of a target group of 336, defined as treated members of the Guidance Study plus the Oakland Growth Study; Stewart and Livson studying 165 (35%) out of 460. Livson and Peskin, Tuddenham and Elder all used the Oakland Growth Study only, obtaining data on 30 percent, 34 percent and 43 percent respectively. At recovery, spouse and offspring were evaluated as well as the subject himself. Principal Variables: Time 1 - Characteristics of the family environment (values, democracy, participation) and of the parents (competence, warmth, harmony) for the early studies; children's personality traits, intelligence, school achievement and economic privation for studies of adults. Time 2 - Personality types, intelligence, personality, smoking, occupation, general adjustment, behavior, IQ, body build of subject's offspring. References: Block (1971), Bronson, et al. (1959), Elder (1974), Hunt and Eichorn (1972), Livson and Peskin (1967), MacFarlane (1964), Stewart and Livson (1966), Stewart (1962), Tuddenham (1959).

B. Borland: Hyperactive Boys

Study Type: Catchup prospective. Index Cases: 37 white males were selected from old child guidance clinic records if they met five criteria: (1) symptoms satisfied criteria for the hyperactive syndrome, (2) each had a brother, (3) aged 4 to 11 at referral, (4) IQ of 80+, (5) no physical or medical problems. Control Cases: Same sexed siblings. Substituted brothers-in-law in two cases. Length of Followup Interval: 20 to 25 years. Interval Measurement: None. Data Sources: Clinic records; family physicians; school records, including grades and IQ scores; and interviews. Case Recovery: Interviews with 20 index cases (49%) and 19 controls. Principle Variables: Time 1 - Number of symptoms of hyperactivity, IQ, school achievement. Time 2 - Adult symptoms of hyperactivity, school success, socioeconomic situation, hyperactivity in offspring. Reference: Borland and Heckman (1976).

C. Cambridge-Somerville Youth Study

Study Type: Real time prospective. Index Cases: One boy randomly chosen from each of 253 pairs matched on age (5 to 13), delinquency proneness, family background and home environment. Pool of subjects came from agency and school recommendations of difficult and average boys in the community. Control Cases: The 253 boys not selected as index cases. Length of Followup Interval: 35 years. Interval Measurement: 200 interviewed at about age 23; records searched when they were 31. Data Sources: Police, mental hospitals, alcohol treatment centers, questionnaires. Case Recovery: 477 located out of 506 (94%). Questionnaire information on 222 (44%). Interviewing still in progress. Principal Variables: Time 1 - Family type, child's delinquency, treatment. Time 2 - Alcoholism, adult criminal behavior. References: McCord (1976a and b), McCord and McCord (1959), McCord, et al. (1960), Powers and Witmer (1951).

D1. Danish Adoption Studies: Criminality

Study Type: Catchup prospective. Index Cases: Part 1 - 1145 male adoptees born 1927 to 1941. Part 2 - 143 criminal adoptees whose biological fathers were born since 1890. Control Cases: Part 1 - 1145 nonadoptees matched on age, sex, occupational status of father and residence. Part 2 - 143 adoptees not known to the police, matched for age and adoptive father's occupation. Length of Followup Interval: Birth to age 30 to 44 years. Interval Measurement: None. Data Sources: Adoption records, police records, psychiatric register, midwives' reports. Case Recovery: Information obtained on 971 biological fathers of 1145 adoptees (85%) and on 1120 fathers of 1145 nonadoptees (98%). Principal Variables: Time 1 - Criminal status of biological and adoptive parents. Time 2 - Criminal record of adoptees versus nonadoptees; psychiatric diagnoses of biological parents of criminal and noncriminal adoptees. Reference: Hutchings and Mednick (1975).

D2. Danish Adoption Studies: Schizophrenia

Study Type: Catchup prospective. Index Cases: 173 biological and 74 adoptive relatives of 34 persons adopted between 1924 and 1947 and unanimously classified as schizophrenic by four researchers on the basis of abstracts of mental hospital records (out of 507 hospitalized). Control Cases: 174 biological relatives and 91 adoptive relatives of 34 adoptees who had never lived in a mental institution, matched on age, sex, socioeconomic status of rearing family, time spent with biological relatives, in child care or in foster home before transfer to adoptive home. Length of Followup Interval: From 30 to 50 years.

Interval Measurement: None. Data Sources: Institutional records, interviews. Case Recovery: Interviews with 329 of 512 relatives (64%). Principal Variables: Time 1 - Sex and kinship to index or control case. Time 2 - Psychiatric diagnosis. Reference: Kety, et al. (1968, 1975).

E. Fels Research Institute Study

Study Type: Real time prospective study. Index Cases: 300 newborn, 6 to 15 added each year since 1929. Length of Followup Interval: 19 to 29 years. Interval Measurement: Frequent assessment through age 10. Semi-annual home visits and observations in school and camp settings; biannual interviews in adolescence. Data Sources: Behavioral observations, physical growth measurements, IQ and achievement tests, projective tests, parent ratings, interviews with mother and child. Case Recovery: Kagan and Moss (1960) studied 54 subjects born between 1930 and 1939. 36 percent of those entering the sample in those years. Kagan and Freeman studied 50 subjects. Moss and Kagan (1961) studied 71 subjects aged 20 to 29. Principal Variables: Time 1 - IQ, socioeconomic status, early personality traits, maternal rearing patterns. Time 2 - adolescent and adult personality (achievement, aggression, sexuality, passivity, dependence, anxiety, etc.) and intelligence. References: Kagan and Freeman (1963), Kagan and Moss (1960, 1962), Moss and Kagan (1961).

F. Forssman and Thuwe: Refused Abortions

Study Type: Catchup prospective. Index Cases: 120 children surviving to age 21 out of pregnancies which went to term after the mother requested and was refused an abortion. Applications for abortion were during 1939 to 1941 in Goteborg. Control Cases: The next same sexed child born in the same hospital and surviving (N = 119). Length of Followup Interval: Birth to age 21. Interval Measurement: None. Data Sources: Registry offices, child welfare boards, school records, hospitals, guidance clinics, mental hospitals, penal register, military, temperance boards. Case Recovery: All were located in records. Principal Variables: Time 1 - Maternal age, socioeconomic status, broken home. Time 2 - Psychiatric treatment, alcoholism, education, military service, delinquency, marital status, receipt of public assistance. Reference: Forssman and Thuwe (1966).

G. Glueck: Delinquent Boys:

Study Type: Real time prospective. Index Cases: 500 white males committed to a correctional school. Control Cases: 500 schoolboys free of both official and unofficial delinquency matched by residence, age, ethnicity, global IQ. Length of Followup Interval: 15 to 22 years; boys aged 9 to 17 at intake were followed at age 31. Interval Measurement:

At age 25. Data Sources: Criminal data from police and court sources; interviews with subject, parents, wife, others; records of settlement houses, boys' clubs, family welfare agencies. Case Recovery: 463 delinquents and 466 nondelinquents at age 25 (83%), 438 delinquents and 442 nondelinquents at age 31 (88%). Principal Variables: Time 1 - Seriousness of offense history. Time 2 - Educational attainment, marital stability, criminality, mobility, health, psychiatric care, occupational attainment, leisure interests, military history. Reference: Glueck and Glueck (1968).

H.   Havighurst: River City

Study Type: Real time followup. Index Cases: 487 children aged 11. Length of Followup Interval: Nine years (to age 20). Interval Measurement: Tested from sixth through twelfth grades. Data Sources: Intelligence tests; ratings by teachers; sociometric tests; talent tests; psychological tests; interviews with subject, employers, college counselors; court records. Case Recovery: Most of the analysis of adult adjustment are limited to 411 individuals (84%). Principal Variables: Time 1 - IQ, leadership, talent, personal and social adjustment, social class. Time 2 - School progress, including college attendance; early marriage; work adjustment; competence (work history, adjustment in college, marital history, criminal behavior). Reference: Havighurst, et al. (1962).

I.   Heber: Children of Mothers with Low IQs

Study Type: Real time prospective. Index Cases: Half of 40 infants born to mothers with IQs less than 70 in an urban poverty area, randomly assigned to treatment of mother and child. Control Cases: a) half of 40 infants born to mothers with IQs less than 70 randomly assigned to control status, b) siblings of treatment and control cases. Length of Followup Interval: Birth to age 7. Interval Measurement: Very frequent (every three weeks). Data Sources: IQ tests, tests of language development. Case Recovery: All recovered. Principal Variables: Time 1 - Mother's IQ, infant's IQ. Time 2 - Children's IQ and language development, mother-child interaction, speech patterns. References: Heber (1971), Heber and Garber (1974).

J.   Jessors: Colorado Students

Study Type: Real time prospective. Index Cases: A random stratified sample of 1126 students from three junior high schools, 2290 from three high schools and 497 from a college. Length of Followup Interval: Four years for most items. Interval Measurement: Yearly questionnaire administration. Data Sources: Questionnaires given to subjects in groups. Case Recovery: Four years' data available on 432 (30%) of junior high sample, 205 (41%) of college sample. Two years' information available on 692 (31%) of the students in the high

school sample. Principal Variables: Time 1 - Personal values; perceptions; beliefs; conventionality; history of drinking, sexual experience, marijuana use; deviance; church attendance; grade point average. Time 2 - Year of transition to drinking, marijuana use, nonvirginity; attitudes toward parents and peers. References: Jessor (1976), Jessor and Jessor (1975), Jessor and Jessor (1975), Jessor, et al. (1973).

K.   Kandel: New York State High School Students

Study Type: Real time prospective. Index Cases: 10130 New York State public secondary school students. Length of Followup Interval: five to six months. Interval Measurement: None. Data Sources: Questionnaires to students and parents. Case Recovery: 5468 (54%) with Time 2 questionnaires that could be matched to Time 1 questionnaire. Principal Variables: Time 1 - Amount and type of drug use, parents' drug use, best friend's drug use, attitudes toward drug use, year in high school, ethnicity. Time 2 - Amount and type of drug use. References: Kandel (1975a and b), Kandel, et al. (1975, 1976).

L.   Lane and Albee: IQs of Schizophrenics

Study Type: Follow back. Index Cases: 36 schizophrenics in state and Veterans Administration hospitals who had no sign of childhood schizophrenia and had school test information available for themselves and a sibling; spouses and children of schizophrenics who attended the same schools. Control Cases: 35 cases matched on IQ score in second grade (age seven) identified in school records. Control had to have a sibling with an available IQ score. Length of Followup Measurement: Second grade to adult. Maximum interval approximately 24 years. Interval Measurement: None. Data Sources: School and hospital records. Case Recovery: Unknown. Cases admitted only if recoverable. Principal Variables: Time 1 - IQ scores of schizophrenics, matched controls and siblings in second grade. Time 2 - Adult psychiatric status, IQs in records of the same school district for children and spouses of schizophrenics. References: Lane and Albee (1964, 1965, 1968), Lane, et al. (1970).

M.   Langner: Manhattan Child Followup

Study Type: Real time prospective. Index Cases: 1034 children aged 6 to 18, a representative sample of Manhattan (New York City) children. Length of Followup Interval: Five years. Interval Measurement: None. Data Sources: Interviews with mothers and children. Case Recovery: 732 cases (71%). Principal Variables: Time 1 - Stressful life events (illness, change in economic well being in the household); behavior patterns (dependence, fighting, conflict, isolation, weak group membership); family functioning; race. Time 2 - Disturbed behavior patterns; behavioral stability and change. References: Gersten, et al. (1976), Langner, et al. (1976).

N.   Mednick and Schulsinger: Children of Schiziphrenic Mothers

     Study Type: Real time prospective. Index Cases: 207 children of
schizophrenic mothers, with median age of 15. Control Cases: 104
subjects with no known mental illness in parents or grandparents matched
for age, sex, social class, years of education, placement out of parental
home, rural-urban residence. Length of Followup Measurement: 8 to 12
years (so far). Followup at ages 18 to 30. Interval Measurement: Inter-
view and check of psychiatric register after four years. Data Sources:
Clinical interviews; psychophysiological assessments; psychiatric and
psychological exainations; interviews on life history and school
behavior; parent interviews, school reports, midwife's report; check
of psychiatric register for child, parents and other relatives; mother's
hospital records. Case Recovery: 173 of 207 high risk subjects; 91
of 104 low risk subjects. Principal Variables: Time 1 - Psychophysio-
logical status, psychiatric problems, mother's premorbid history,
severity of mother's illness, contact with mother and father, perinatal
complications. Time 2 - Current diagnostic status. References: Mednick
(1973), Mednick (1970), Schulsinger (1976).

O.   Miller and Dyer: Preschool Programs

     Study Type: Real time prospective. Index Cases: From the pool
of those who registered for Head Start classes in four areas of Louisville,
Kentucky, students were randomly selected and assigned to one of four
experimental Head Start nursery school programs: Bereiter-Engelmann;
Darcee; Montessori; and Traditional. By the end of the prekindergarten
years the total in experimental classes was 214. Control Cases: Two
control samples: (1) 34 children, 22 drawn from waiting lists for Head
Start and 12 from names supplied by the teachers and principals, often
preschool brothers and sisters of children enrolled in the school;
(2) 15 low income children who had not attended Head Start but had
entered special kindergarten classes using behavior modification
principles. Length of Followup Interval: Four years, except for the
second control group which had three years. Interval Measurement:
Assessment of the four treatment modalities to demonstrate their
difference; psychological tests, behavioral inventories. Data Sources:
Tests and rating scales on subjects; information forms filled out by
parents. Case Recovery: 175 of 214 experimental cases; 29 of 34
Control Group 1 cases; 11 of 15 Control Group 2 cases. Principal
Variables: Time 1 - Prekindergarten IQ, program type, sex. Time 2 -
Achievement, IQ, behavior. Reference: Miller and Dyer (1975).

P.   National Child Development Study

     Study Type: Catchup prospective. Index Cases: 17418 children
comprising a one week birth cohort in England, Wales and Scotland in
March 1958, originally studied in a survey of the causes of perinatal

death. Length of Followup Interval: Birth to eight years. Interval
Measurement: These are interval studies. Children were followed at
age seven. There is a still unpublished followup at age eleven. Data
Sources: At birth: obstetric, sociological and medical information.
At age seven: teachers' reports, achievement and psychological tests,
medical examination, interviews with biological or adoptive parents,
records of adoption. Case Recovery: 15468 in followup (89%). For
study of illegitimate births, 526 of 679 followed up (77%). For study
of adopted children, 145 of 205 interviews with adoptive parents (71%).
Principal Variables: Time 1 - Obstetric factors, illegitimacy, adoptive
status, sex, class, home characteristics. Time 2 - School achievement,
adjustment, intelligence, physical development. References: Crellin,
et al. (1971), Davie, et al. (1972), Seglow, et al. (1972).

Q.   The National Survey of Health and Development

     Study Type: Real time prospective. Index Cases: A birth cohort
of 5362 individuals born in Great Britain for the first week of March
1946. Control Cases: A subgroup was excluded from tests to provide a
comparison with those given repeated evaluations. Length of Followup
Interval: Birth to thirties. Interval Measurement: IQ tests at ages
8, 11 and 15. Teacher ratings at ages 13 and 15. Self rating psycho-
logical inventory at 13. School physical examinations at ages 6, 7,
11 and 15. Parent questionnaires at ages 4, 6, 8, 9, 11 and 15. Data
Sources: Teachers, parents, doctors, employment offices, police and
court records, school achievement records, examinations, questionnaires
completed by survey member, school setting descriptions from school.
Case Recovery: 77 percent with full educational test results; 98 percent
have some followup information. Principal Variables: Time 1 - Home
environment (insecurity, family breakup, education encouragement), social
class, early symptoms of disturbed behavior, hospital admissions. Time
2 - School attainment, delinquency, enuresis, illegitimacy. References:
Douglas (1964, 1966, 1970, 1975), Douglas, et al. (1968), Wadsworth (1976).

R1.   Robins: St. Louis Child Guidance Study

     Study Type: Catchup prospective. Index Cases: 503 consecutive
white patients with an IQ over 80 seen at a child guidance clinic who
survived past age 25. Control Cases: 91 subjects with problem free
elementary school records matched on race, age, sex, IQ and socio-
economic status who survived past age 25. Length of Followup Interval:
Median age at referral was 13. Interviewed approximately 30 years later.
Interval Measurement: None. Data Sources: Clinic records, police,
school, hospital and many other records. Personal interview. School
and police records for their sons. Case Recovery: Interviews obtained
for 82 percent; some records for 98 percent. Principal Variables:

Time 1 - Symptoms in childhood; antisocial and other behavior problems of parents; childhood social status. Time 2 - Psychiatric diagnosis (in particular, sociopathy), crime, alcoholism, marital stability, mortality, school and police records of offspring. References: Robins (1974), Robins and O'Neal (1958a and b), Robins and Lewis (1966), Robins, et al. (1962a and b).

R2. Robins: Black School Boys as Adults

Study Type: Catchup prospective. Index Cases: 235 black males born and raised in St. Louis, Missouri with an IQ of 85 or higher, selected from elementary school records. Length of Followup Interval: Approximately 20 years. Interval Measurement: None. Data Sources: School, police, military and other record sources for men, wives and children. Interview with subject. Case Recovery: Interviews for 223; records for all. Principal Variables: Time 1 - Home environment, early childhood behavior, school success, juvenile delinquency. Time 2 - Adult psychiatric status, adult deviance (crime, drug abuse, alcoholism), educational and economic outcomes, school success and delinquency of offspring. References: Robins (1972), Robins and Hill (1966), Robins and Murphy (1967), Robins and Wish (1977), Robins, et al. (1968, 1971, 1975).

S. Roff: Military Service of Child Guidance Cases

Study Type: A. Catchup prospective; B. Follow back. Index Cases: Cases for which adequate information is available both in military and child guidance clinic records. For catchup study: 265 delinquents; for follow back study: 104 with military discharge for psychoneurosis and 164 with discharge for bad conduct. Control Cases: For catchup, randomly chosen school boys. For follow back, high ranking enlisted men without discipline problems who had been clinic patients. Length of Followup Interval: Approximately eight years. Interval Measurement: None. Data Sources: Child guidance clinic records; military acceptance or rejection, military discharge; Veterans Administration records; school records. Case Recovery: Not reported. Principal Varibles: Time 1 - Childhood maladjustment, personality problems, family background (status, interaction, structure), delinquency, peer relations. Time 2 - Success in military service. References: Roff (1960, 1969, 1970, 1972).

T1. Rutter: Isle of Wight Study

Study Type: Real time followup. Index Cases: Isle of Wight. A total school population of 10 to 11-year-olds (N = 2234). Control Cases: A sample of all 10-year-olds in an inner London borough. Length of Followup Interval: Four years. Interval Measurement: None on the total sample. Some ongoing intervention studies during the interval. Data Sources: Tests, teacher and parent questionnaires, medical exams,

interviews with children, survey questionnaires to subjects at Time 2. Case Recovery: Not clear. Principal Variables: Time 1 - Family characteristics; nature and severity of psychiatric, physical and intellectual-educational problems. Time 2 - Continued manifestation of problem behaviors. References: Rutter, et al. (1970, 1975, 1976a and b).

T2. Rutter and Lockyer: Autistic Children

Study Type: Catchup prospective. Index Cases: 64 prepubescent patients with a hospital diagnosis of childhood psychosis or autism. Control Cases: 63 nonpsychotic psychiatric patients of the same sex and IQ seen the same year. Length of Followup Interval: 12 to 20 years. Interval Measurement: at 9 to 10 years. Data Sources: Hospital records (Time 1); neurological and psychiatric examination, observation, interview with parent, psychological tests (interval evaluation); mail questionnaires to parents and hospital and clinic records (final assessment). Case Recovery: All psychotics followed; 61 controls included in the interval evaluation; not followed for the final assessment. Principal Variables: Time 1 - IQ, presence or absence of speech, neurological signs, response to sound, behavior problems, interpersonal relationships. Time 2 - Institutionalization, educational level, marriage, occupation, speech, social skills, convulsions, IQ, psychological tests. References: Lockyer and Rutter (1969), Rutter (1970), Rutter, et al. (1967).

U. Terman Study of the Gifted

Study Type: Real time prospective. Index Cases: 1528 children in grades 1 to 12 in California schools who scored in the top 1 percent on an IQ test. Control Cases: 533 unselected subjects given psychological tests. Length of Followup Interval: Forty years. Interval Measurement: Frequent: 3 field studies at 6, 18 and 30 years after entry. Annual reporting by mail in first five years. Mail contact after 14 years, 23 years, 28 years and 33 years. Data Sources: Interviews with family and subject, questionnaires to parents, teachers, subjects; teachers' observations, medical examinations, school records, school achievement tests, home visits. Case Recovery: 1188 replied to 40-year-later Information Blank (74% of total; 85% of survivors); only 2 percent entirely lost. Principal Variables: Time 1 - Family values, personality traits and motivation, health and school progress. Time 2 - Mortality, health, psychological adjustment, alcohol use, crime homosexuality, occupation, marital status, fertility, income. References: Oden (1968), Terman and Oden (1947, 1959).

V.   Thomas and Chess: New York Longitudinal Study

Study Type: Real time prospective. Index Cases: 136 children followed from infancy. Length of Followup Interval: 16 years (for the 95 older cases). Interval Measurement: Frequent during first five years. Data Sources: Interviews with parent and child; clinical psychiatric evaluation; behavioral and psychometric data. Case Recovery: 90 of 95 who had reached age 16. Principal Variables: Time 1 - Early personality traits ("temperament"), parent attitudes; early childhood behavior problems. Time 2 - New behavioral disturbance in adolescence, remission. References: Thomas and Chess (1976), Thomas, et al. (1968).

W.   Watt: School Records of Schizophrenics

Study Type: Follow back. Index Cases: 54 schizophrenics aged 15 to 34 at admission to Massachusetts mental hospitals between 1958 and 1965 who were found to have school records in one town. Control Cases: 143 cases drawn from the school records and matched for age, sex, race, father's occupation and father's education, who had never been in a mental hospital in Massachusetts. Three controls were chosen for most index cases. Length of Followup Interval: 10 to 28 years (from school entry to admission to hospital). Interval Measurement: None. Data Sources: Hospital register, school records. Case Recovery: Selection required presence in both Time 1 and Time 2 records. Principal Variables: Time 1 - Behavior patterns in school (e.g., conscientiousness, assertiveness, conformity), school achievement, IQ. Time 2 - Adult diagnosis of schizophrenia. References: Watt, et al. (1970, 1976).

X.   West and Farrington: Cambridge Study in Delinquent Development

Study Type: Real time prospective. Index Cases: 411 boys aged 8 to 9 chosen from two successive classroom generations in the schools in a London working class area. Length of Followup Measurement: 14 years (continuing). Interval Measurement: Tested at age 10 and 14; interviewed at age 16 and 18. Data Sources: Police records for subject, parents and siblings; medical and social service records; teachers' questionnaires; peer ratings; tests; parent interviews. Case Recovery: 389 boys at age 19. Principal Variables: Time 1 - Income, family size, parent and sibling criminality, quality of parental behavior, child's aggressiveness. Time 2 - Delinquency. References: Farrington, et al. (1975), West and Farrington (1973).

Y.   Wolfgang: Delinquency in a Cohort

Study Type: Catchup prospective. Index Cases: 9945 boys born in 1945 who lived continuously in Philadelphia from ages 10 to 18, selected from public, parochial and private schools. Length of Followup Interval:

Birth to end of juvenile court jurisdiction (age 18). Interval Measurement: None. Data Sources: School and police records. Case Recovery: Not an issue. Principal Variables: Time 1 - Race, school attended, IQ, achievement level, final grade reached, census tract of last residence (used for social class assignment), physical handicap, residential moves. Time 2 - Severity of offenses, recidivism, type of crime. Reference: Wolfgang, et al. (1972).

Z.   Youth in Transition

Study Type: Real time prospective. Index Cases: 2281 10th-grade boys (age 15) selected in a multistage national sample (1966). Length of Interval Covered: Eight years. Interval Measurement: Recontacted after 18 months, 30 months and 42 months. Data Sources: Personal interview, tests, questionnaires; questionnaires for teachers, principals, school counselors. Case Recovery: 1365 found for fifth wave (60%). Principal Variables: Time 1 - Home environment, family size, family values, residence, race, socioeconomic status, academic progress, grades, delinquency, IQ, self concept, motivation, values. Time 2 - High school dropout, values and attitudes in high school, enlisting in the military, drug use, smoking, alcohol, delinquency, employment. References: Bachman (1970), Bachman, et al. (1971), Johnston (1973), Johnston and Bachman (1972).

References

Achenbach, T.M. Developmental psychopathology. New York: Ronald, 1974, Chapter II.

Annell, A.L. The prognosis of psychotic syndromes in children. Acta psychiatr. scand. 39 (Supple. 172), 235-297 (1963).

Bachman, J.G. Youth in transition. Vol. II: The impact of family background and intelligence on tenth-grade boys. Ann Arbor: Institute for Social Research, 1970.

Bachman, J.G., Green, S., Wirtanen, I.D. Youth in transition. Vol. III: Dropping out - Problem or symptom. Ann Arbor: Institute for Social Research, 1971.

Baltes, P.B. Longitudinal and cross-sectional sequences in the study of age and generation effects. Human Development, 11, 145-171 (1968).

Baltes, P.B., Schaie, K.W. & Nardi, A.H. Age and experimental mortality in a seven-year longitudinal study of cognitive behavior. Developmental Psychology, 6, (1), 78-84 (o972).

Barker, D.J.P & Edwards, J.H. Obstetric complications and school performance. Brit. Journal of Medicine, 3, 695-699 (1967).

Bender, L. The life course of children with schizophrenia. American Journal of Psychiatry, 130, 783-786 (1973).

Block, J. Lives through time. Berkeley: Bancroft Books, 1971.

Bohman, M. A study of adopted children, their background, environment and adjustment. Acta paediat. scand., 61, 90-97 (1972).

Borland, B.L., Heckman, H.K. Hyperactive boys and their brothers. Arch. gen. Psychiat., 33, 669-675 (1976).

Bronson, W.C., Katten, E.S. & Livson, N. Patterns of authority and affection in two generations. Journal of Abnormal Social Psychology, 58, 143-152 (1959).

Campbell, D.T. & Boruch, R.F. Making the case for randomized assignment to treatments by considering the alternatives. In C.A. Bennett and A.A. Lumsdaine (Eds.), Evaluation and experiment. New York: Academic Press, 1975.

Campbell, D.T. & Erlebacher, A.E. How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In J. Hellmuth (Ed.) Compensatory education: A national debate, Vol. 3. New York: Brunner Mazel, 1970.

Cooper, J.E. The use of a procedure for standardizing psychiatric diagnosis. In E.H. Hare and J.K. Wing (Eds.), Psychiatric epidemiology. Oxford: Oxford University Press, 1970.

Crellin, E., Kellmer Pringle, M.D. & West, P. Born illegitimate. London: National Foundation for Educational Research in England and Wales, 1971.

Davie, R., Butler, N. & Goldstein, H. From birth to seven. London: Longman, 1972.

Douglas, J.W.B. The home and the school. London: Macgibbon and Kee, 1964.

Douglas, J.W.B. The school progress of nervous and troublesome children. British Journal of Psychiatry, 112, 1115-1116 (1966).

Douglas, J.W.B. Broken families and child behavior. J. roy. coll. physicians Lond., 4, 203-210 (1970).

Douglas, J.W.B. Early hospital admissions and later disturbances of behavior and learning. Develop. Med. Child Neurol., 17, 456-480 (1975).

Douglas, J.W.B., Ross, J.M. & Simpson, H.R. All our future. London: Peter Davies, 1968.

Eaton, L. & Menolascino, F.J. Psychotic reactions of childhood: A followup study. Amer. J. Ortho-psychiat., 37, 521-529 (1967).

Eisenberg, L. The course of childhood schizophrenia. Arch. Neurol. Psychiat., 78, 69-83 (1957).

Elder, G.H. Children of the great depression. Chicago: University of Chicago Press, 1974.

Farrington, D.P., Gundry, G. & West, D.J. The familial transmission of criminality. Med. Sci. Law, 15, 177-186 (1975)

Feinstein, A.R. Clinical judgment. Baltimore, Md.: Williams and Wilkins, 1967.

Forssman, H. & Thuwe, I. One hundred and twenty children born after application for therapeutic abortion refused. Acta psychiat. scand., 42, 71-88 (1966).

Gersten, J.C., Langner, T.S., Eisenberg, J.G., Simcha-Fagan, O. & McCarthy, E.D. Stability and change in types of behavioral disturbance of children and adolescents. Journal of Abnormal Child Psychology, 4, 111-127 (1976)

Gittelman-Klein, R. & Klein, D.F. Methylphenidate effects in learning disabilities. Arch, gen. psychiat., 33, 655-664 (1976).

Glueck, S & Glueck, E. Juvenile delinquents grown up. New York: The Commonwealth Fund, 1940.

Glueck, S. & Glueck, E. Delinquents and nondelinquents in perspective. Cambridge: Harvard University Press, 1968.

Goodman, L.A. The multivariate analysis of qualitative data: Interaction among multiple classifications. J. Amer. Statist. Assoc., 65, 225-256 (1970).

Hagnell, O. A prospective study of the incidence of mental disorder. Stockholm: Scandinavian University Books, 1966.

Hanson, D.R., Gottesman, II. & Heston, L.L. Some possible childhood indicators of adult schizophrenia inferred from children of schiophrenics. British Journal of Psychiatry, 129, 142-154 (1976).

Havighurst, R.J., Bowman, P.H., Liddle, G., Matthews, C. & Pierce J. Growing up in River City. New York: Wiley, 1962.

Heber, R. Rehabilitation of families at risk for mental retardation: A progress report. Madison, Wisc.: Rehabilitation Research and Retraining Center in Mental Retardation, 1971.

Heber, R. & Garber, H. Progress Report III: An experiment in the prevention of cultural-familial retardation. In Proceedings of the Third International Congress of the International Association for the Scientific Study of Mental Deficiency (1974).

Heise, D.R. Causal inference from panel data. In E.F. Borgatta and G.W. Bohrnstedt (eds.), Sociological methodology. San Francisco: Jossey-Bass, 1970.

Hunt, J.V. & Eichorn, D.H. Maternal and child behaviors: A review of data from the Berkeley Growth Study. Semin. Psychiat., 4, 367-381 (1972)

Hutchings, B. & Mednick, S.A. Registered criminality in the adoptive and biological parents of registered male criminal adoptees. In R. Fieve, D. Rosenthal & H. Brill (Eds.), Genetic research in psychiatry. Baltimore, Md.: Johns Hopkins, 1975.

Jessor, R. Predicting time of onset of marijuana use: A developmental study of high school youth. J. consult. clin. psycholo., 44, 1:125-134 (1976).

Jessor, R. & Jessor, S.L. Adolescent development and the onset of drinking: A longitudinal study. J. stud. alcohol, 36, 1:27-51 (1975).

Jessor, S.L. & Jessor, R. Transition from virginity to nonvirginity among youth: A social-psychological study over time. Developmental Psychology, 11, 4:473-484 (1975).

Jessor, R., Jessor, S.L. & Finney, J. A social psychology of marijuana use: Longitudinal studies of high school and college youth. J. personal soc. psychol., 26, 1:1-15 (1973).

Johnston, L. Drugs and American youth. Ann Arbor: Institute for Social Research, 1973.

Johnston, J. & Bachman, J.G. Youth in transition. Vol. V: Young men and military services. Ann Arbor: Institute for Social Research, 1972.

Joreskog, K.G. & Sorbom, D. Statistical models and methods for analysis of longitudinal data. In D.J. Aigner and A.S. Goldberger (Eds.), Latent variables in socioeconomic models. New York: North-Holland, 1976.

Kagan, J. & Freeman, M. Relation of childhood intelligence, maternal behaviors and social class to behavior during adolescence. Child Develop., 34, 899-911 (1963).

Kagan, J. & Moss, H.A. The stability of passive and dependent behavior from childhood through adulthood. Child Develop., 31, 577-591 (1960).

Kagan, J. & Moss, H.A. Birth to maturity. New York: Wiley, 1962.

Kandel, D. reaching the hard-to-reach: Illicit drug use among high school absentees. Addictive Diseases, 1, 4:465-480 (1975a).

Kandel, D., Single, E. & Kessler, R. The epidemiology of drug use among New York State high school students: Distribution, trends, and change in rates of use. Amer. j. publ. health, 66, 43-53 (1976).

Kandel, D. Stages in adolescent involvement in drug use. Science, 190, 912-914 (1975b).

Kety, S.S., Rosenthal, D., Wender, P.H. & Schulsinger, F. The types and prevalence of mental illness in the biological and adoptive families of adopted schizophrenics. In D. Rosenthal and S.S. Kety (Eds.), The transmission of schizophrenia. Oxford: Pergamon, 1968, pp. 345-362.

Kety, S.S., Rosenthal, D., Wender, P.H., Schulsinger, F. & Jacobsen, B. Mental illness in the biological and adoptive families of adopted individuals who have become schizophrenic: A preliminary report based on psychiatric interviews. In R. Fieve, D. Rosenthal and H. Brill (Eds.), Genetic research in psychiatry. Baltimore, Md.: Johns Hopkins, 1975, pp. 147-165.

Lane, E.A. & Albee, G.W. Early childhood intellectual differences between schizophrenic adults and their siblings. J. abnorm. soc. psychol., 68, 193-195 (1964).

Lane, E.A. & Albee, G.W. Childhood intellectual differences between schizophrenic adults and their siblings. Amer. j. orthophyschiat., 35, 4:747-753 (1965).

Lane, E.A. & Albee, G.W. On childhood intellectual decline of adult schizophrenics: A reassessment of an earlier study. J. abnorm. psychol., 73, 2:174-177 (1968).

Lane, E.A., Albee, G.W. & Doll, L.S. The intelligence of children of schizophrenics. Developmental Psychology, 2, 3:315-317 (1970).

Langner, T.S., Gersten, J.C., McCarthy, E.D., Eisenberg, J.C., Greene, E.L., Herson, J.H. & Jameson, J.D. A screening inventory for assessing psychiatric impairment to children 6 to 18. J. consult. clin. psychol., 44, 286-296 (1976).

Livson, N. & Peskin, H. Prediction of adult psychological health in a longitudinal study. J. abnorm. psychol., 72, 509-518 (1967).

Lockyer, L. & Rutter, M. A five to fifteen-year followup study of infantile psychosis III: Psychological aspects. Brit. J. Psychiat., 115, 865-882 (1969).

McCord, J. Patterns of deviance. Paper presented at the Annual Meeting of the Society for Life History Research in Psychopathology, 1976a.

McCord, J. A thirty-year followup of treatment effects. Paper presented at the Meetings of the American Association of Psychiatric Services for Children, 1976b.

McCord, W. & McCord, J. Origins of crime. New York: Columbia University Press, 1959.

McCord, W., McCord, J. & Gudeman, J. Origins of alcoholism. Palo Alto: Stanford University Press, 1960.

MacFarlane, J.W. Perspectives on personality and change from the Guidance Study. Vita hum., 7, 115-126 (1964).

Mednick, B.R. Breakdown in high-risk subjects: Familial and early environmental factors. J. abnorm. psychol., 82, 469-475 (1973).

Mednick, S.A. Breakdown in individuals at high risk for schizophrenia: Possible predispositional perinatal factors. Ment. Hygiene, 54, 50-63 (1970).

Mednick, S.A., Mura, E., Schulsinger, F. & Mednick, B. Erratum and further analysis: "Perinatal conditions and infant development in children with schizophrenic parents." Social Biol., 20, 111-112 (1973).

Michael, C.M., Morris, D.P. & Soroker, E. Followup studies of shy, withdrawn children. II: Relative incidence of schizophrenia. Amer. J. Orthopsychiat., 27, 331-337 (1957).

Miller, L.B. & Dyer, J.L. Four preschool programs: Their dimensions and effects. Monogr. Soc. Res. Child Develop., 40, 1-170 (1975).

Moos, H.A. & Kagan, J. Stability of achievement and recognition seeking behaviors from early childhood through adulthood. J. abnorm. soc. psychol., 62, 504-513 (1961).

Niswander, K. & Gordon, M. The women and their pregnancies. Philadelphia: Saunders, 1972.

Oden, M.H. The fulfillment of prmoise: 40 year followup of the Terman gifted group. Genet. Psychol. Monogr., 77, 3-93 (1968).

O'Donnell, J.A., Voss, H.L., Clayton, R.R., Slatin, G.T. & Room, R.G. Young men and drugs: A nationwide survey. Rockville, Md.: NIDA Research Monograph Ser. 5: DHEW Publ. No. ADM-76-311 (1976).

Powers, E. & Witmer, H. An experiment in the prevention of delinquency. New York: Columbia University Press, 1951.

Record, R.G., McKeown, T. & Edwards, J.H. The relation of measured intelligence to birth order and maternal age. Ann. hum. genet. Lond., 33, 61-69 (1969).

Robins, L.N. An actuarial evaluation of the causes and consequences of deviant behavior in young black men. In M. Roff, L.N. Robins and M. Pollack (Eds.), Life history research in psychopathology, Vol. 2. Minneapolis: University of Minnesota Press, 1972, pp. 137-154.

Robins, L.N. Evaluation of psychiatric services for children in the United States. In J.K. Wing and H. Hafner (Eds.), Roots of Evaluation. London: Oxford University Press, 1973.

Robins, L.N. Deviant children grown up. Huntington, N.Y.: Krieger, 1974.

Robins, L.N., Bates, W.M. & O'Neal, P. Adult drinking patterns of former problem children. In D.J. Pittman and C.R. Snyder (Eds.), Society, culture and drinking patterns. New York: Wiley, 1962a.

Robins, L.N., Gyman, H. & O'Neal, P. The interaction of social class and deviant behavior. Amer. sociol. rev., 27, 480-492 (1962b).

Robins, L.N. & Hill, S.Y. Assessing the contributions of family structure, class and peer groups to juvenile delinquency. J. crim. law criminol. police sci., 57, 325-334 (1966).

Robins, L.N. & Lewis, R.G. The role of the antisocial family in school completion and delinquency: A three-generation study. Sociol. quart., 7, 4:500-514 (1966)

Robins, L.N. & Murphy, G.E. Drug use in a normal population of young Negro men. Amer. j. publ. health, 57, 1580-1596 (1967).

Robins, L.N., Murphy, G.E. & Breckenridge, M.B. Drinking behavior of young urban Negro men. Quart. j. stud. alc., 29, 3:657-684 (1968).

Robins, L.N., Murphy, G.E., Woodruff, R.A., Jr., & King L.J. The adult psychiatric status of black school boys. Arch. gen. psychiat., 24, 338-345 (1971).

Robins, L.N. & O'Neal, P. The marital history of former problem children. Social problems, 5, 347-358 (1958a).

Robins, L.N. & O'Neal, P. Mortality, mobility and crime: Problem children thirty years later. Amer. sociol. rev., 23, 162-171 (1958b).

Robins, L.N., West, P.A. & Herjanic, B. Arrest and delinquency in two generations: A study of black urban families and their children. J. child psychol. psychiat., 16, 125-140 (1975).

Robins, L.N. & Wish, E. Childhood deviance as a developmental process: A study of 223 urban black men from birth to 18. Social forces, 56, 448-473 (1977).

Roff, M. Relation between certain preservice factors and psychoneurosis during military duty. U.S. armed forces med. j., 11, 152-160 (1960).

Roff, M. Juvenile delinquency and military service. In R.W. Little (Ed.), Selective service and American society. New York: Russell Sage, 1969.

Roff, M. Some life history factors in relation to various types of adult maladjustment. In M. Roff and D.F. Ricks (Eds.), Life history research in psychopathology. Minneapolis: University of Minnesota Press, 1970.

Roff, M. A two-factor approach to juvenile delinquency and the later histories of juvenile delinquents. In M. Roff, L.N. Robins and M. Pollack (Eds.), Life history research in psychopathology, Vol. 2. Minneapolis: University of Minnesota Press, 1972.

Rosenberg, M. Test factor standardization as a method of interpretation. Social forces, 41, 53-61 (1962).

Rutter, M. Autistic children: Infancy to adulthood. Semin. psychiat., 2, 435-450 (1970).

Rutter, M., Graham, P., Chadwick, O. & Yule, W. Adolescent turmoil: Fact or fiction? J. child psychol. psychiat., 17, 35-56 (1976b).

Rutter, M.L., Greenfield, D. & Lockyer, L. A five to fifteen year followup study of infantile psychosis: II. Social and behavioral outcome. Brit. j. psychiat., 113, 1183-1199 (1967).

Rutter, M., Tizard, J. & Whitmore, K. Education, health and behavior. London: Longman, 1970.

Rutter, M.L., Tizard, J., Yule, W., Graham, P. & Whitmore, K. Research report: Isle of Wight studies, 1964-1974. Psychol. med., 6, 313-332 (1976a).

Rutter, M.L., Yule, B., Quinton, D., Rowlands, O., Yule, W. & Berger, M. Attainment and adjustment in two geographical areas. III: Some factors accounting for area differences. Brit. j. psychiat., 126, 520-533 (1975).

Schaie, K.W. A reinterpretation of age related changes in cognitive structure and functioning. In L.R. Goulet and P.B. Baltes (Eds.), Life-span developmental psychology. New York: Academic Press, 1970.

Schulsinger, H. A ten-year followup of children of schizophrenic mothers. Acta psychiat. scand., 53, 371-386 (1976).

Seglow, J., Kellmer Pringle, M.L. & Wedge, P. Growing up adopted. London: National Foundation for Educational Research in England and Wales, 1972.

Single, E., Kandel, D. & Johnson, B.D. The reliability and validity of drug use responses in a large scale longitudinal survey. J. drug issues, 5, 426-443 (1975).

Stewart, L. & Livson, N. Smoking and rebelliousness: A longitudinal study from childhood to maturity. J. consult. psychol., 30, 225-229 (1966).

Stewart, L.H. Social and emotional adjustment during adolescence as related to the development of psychosomatic illness in adulthood. Genet. psychol. monogr., 65, 175-215 (1962).

Sudman, S. & Bradburn, N.M. Response effects in surveys. Chicago: Aldine, 1974.

Terman, L.M. & Oden, M.H. The gifted child grows up. Stanford: Stanford University Press, 1947.

Terman, L.M. & Oden, M.H. The gifted group at mid-life. Genetic Studies of Genius, Vol. 5. Stanford: Stanford University Press, 1959.

Thomas, A. & Chess, S. Evolution of behavior disorders into adolescence. Amer. j. psychiat., 133, 539-542 (1976).

Thomas, A., Chess, S. & Birch, H.G. Temperament and behavior disorders in children. New York; New York University Press, 1968.

Tietze, C. Fertility after discontinuation of intrauterine and oral contraception. Internal j. of fertility, 13, 385-389 (1968).

Tuddenham, R.D. & Read, D. The constancy of personality ratings over two decades. Genet. psychol. monogr., 60, 3-29 (1959).

Wadsworth, M.E.M. Delinquency, pulse rates and early emotional deprivation. Brit. j. criminol., 16, 245-256 (1976).

Watt, N.F. & Lubensky, A.W. Childhood roots of schizophrenia. J. consult. clin. psychol., 44, 363-375 (1976).

Watt, N.F., Stolorow, R.D., Lubensky, A.W. & McClelland, D.C. School adjustment and behavior of children hospitalized for schizophrenia as adults. Amer. j. orthopsychiat., 40, 637-657 (1970).

West, D.J. & Farrington, D.P. Who becomes delinquent? London: Heinemann, 1973.

Wilson, H. Delinquency and child neglect. London: Allen and Unwin, 1962.

Wolfgang, M.E., Figlio, R.M. & Sellin, T. Delinquency in a birth cohort. Chicago: University of Chicago Press, 1972.

Zajonc, R.B. Family configuration and intelligence. Science, 192, 227-236 (1976).

Chapter 5

Sequential Strategies and the Role of Cohort Effects

in Behavioral Development: Adolescent Personality

(1970-1972) as a Sample Case**

John R. Nesselroade and Paul B. Baltes

The Pennsylvania State University

Sequential Strategies and the Role of Cohort Effects

in Behavioral Development: Adolescent Personality

(1970-1972) as a Sample Case

John R. Nesselroade and Paul B. Baltes

The Pennsylvania State University

Introduction

Background

Discrepant findings resulting from the application of longitudinal and cross-sectional designs (e.g., Baltes, 1968; Damon, 1965; Kuhlen, 1963; Schaie, 1970) concerned developmental psychologists for several decades before a concerted effort was made to reach a methodological resolution. The solution was based on the direct recognition of the impact of historical (secular, generational, cohort, etc.) conditions on ontogenetic (age) development. While in the past most developmental psychologists behaved as if the world "stood still," as if it were invariant, the dialectic (Riegel, 1976) and methodological position (Kuhlen, 1963; Riley, Johnson & Foner, 1972; Ryder, 1965; Schaie, 1965) advanced was that methodologies were needed which jointly recognized processes of ontogeny and cultural change.

In developmental psychology, Schaie's (1965) methodological contributions and substantive work on adult intelligence (Schaie, 1979) were seminal. Capitalizing on a long tradition of related work (see Baltes, 1968; Baltes, Cornelius & Nesselroade, 1978, for reviews), Schaie presented a General Developmental Model in 1965 which incorporated three design components (age, cohort and time of measurement)

in representations of developmental functions. Many subsequent empirical studies in the personality and ability domains have provided strong support for the efficacy of the general approach to describing and understanding behavioral development (see, e.g., Baltes, Cornelius & Nesselroade, 1978, 1979; Nesselroade & Baltes, 1974, for review) and the investigation of generation or cohort differences has become theoretically and substantially compelling. There remains much disagreement about methodological specifics (data analysis, interpretation of effects) of Schaie's General Developmental Model but the basics of the approach continue to survive.

Although the present chapter focuses primarily on cohort effects, a few more general observations on methodology are in order. In conjunction with the methodological issues mentioned above, a number of other methodological problems of cross-sectional vs. longitudinal research have been clarified. Simple developmental designs have a variety of weaknesses in internal and external validity (Baltes, Reese & Nesselroade, 1977). Selection effects in sampling, drop-out and survival of subjects, as well as testing effects and generation or cohort differences affect these designs in various ways. Both Baltes (1968) and Schaie (1965) argued that discrepancies and contradictions between cross-sectional and longitudinal outcomes result from violations of basic design assumptions and weaknesses in controls; cohort effects represent but one such design threat.

As to the role of cohort effects, the designs proposed were combinations of several cross-sectional and longitudinal studies and were labeled sequential strategies. In sociology, the term cohort analysis (Riley et al., 1972; Ryder, 1965) was introduced. It was soon recognized, particularly by Riegel (1976) that the problem was not only one of adequate methodology for the assessment of ontogenetic change. Marked cohort effects have implications for theory as well, for example, for the relative usefulness of distinct developmental models or theories (e.g., mechanistic versus organismic). The presence of cohort effects in developmental data challenges the validity of personological orientations for representing developmental processes. At the same time, substantial cohort effects reinforce the need for cooperative efforts by researchers who hold evolutionary, sociological, anthropological and psychological world views of the nature and etiology of human development. This focus on the interactive relationships between individual and historical change is paralleled by important methodological and theoretical contributions in the field of sociology (e.g., Bengtson & Black, 1973; Elder, 1975, 1979; Keniston, 1970, 1971; Neugarten & Datan, 1973; Riley, Johnson & Foner, 1972; Ryder, 1965). But also in developmental psychology, scholars such as Keniston (1971) and Riegel (1972) in particular, were persuasive in charging that research and theory building are much too culture-centered and historically parochial.

The Role of the Cohort Variable in Developmental Theory

Although the principal focus of the study we are reporting was to identify the magnitude and possible role of cohort effects in adolescent personality at a descriptive level, a proper setting for understanding those outcomes will be provided by some initial consideration of certain aspects of explanatory research and theory building involving cohort effects. The logical status of cohort is similar to that of chronological age (e.g., Baer, 1970 Baltes & Goulet, 1971; Wohlwill, 1973). Defined as birth cohort, it is a person or assigned variable (Kerlinger, 1964) rather than an experimental one. The design status given the cohort variable has several implications for theory construction. Clearly, the cohort variable will be regarded quite differently by researchers depending on the conceptual beliefs they hold concerning the significance of the roles of biocultural change, experimental design and the need for process-oriented approaches to the study of development. In the limited space available here we will consider three alternative conceptions of the nature of the cohort variables which bear on the present study. These ideas are elaborated elsewhere (Baltes, Cornelius & Nesselroade, 1978, 1979).

Cohort as Error or Disturbance

This view of cohort follows from an approach in which establishing average and fairly invariant developmental principles (e.g., age-behavior functions) are one's primary objective. Variations and contingencies of developmental functions are essentially disregarded (entered in the "error" term) in specifying the focal developmental principle. From this perspective, cohort effects and the nature of biological and cultural history may be viewed as providing one such a framework of "irrelevant" contingencies.

Historically, Quetelet (1842) was the first to state the general position. Currently, Wohlwill's (1973; see also McCall, 1977) view of the nature of developmental functions is a good illustration of similar ideas. The objective of developmental research would be to isolate basic developmental-ontogenetic processes against a backdrop of "disturbances," however pronounced the latter may be, created by cohort-related influence patterns. Viewing cohort-related phenomena as errors or disturbance seems most likely to occur among psychologists who are oriented toward child rather than life-span development conceptions and who are interested in either so-called basic processes (learning, cognition, etc.) or organismic, growth-oriented conceptions of development.

Cohort as a Dimension of Generalization

Another treatment of the cohort variable can be illustrated by reference to the concept of external design validity (Baltes, Reese & Nesselroade, 1977; Bell & Hertz, 1976; Campbell & Stanley, 1963). External validity involves making accurate generalizations about relationships from a particular set of data to other potential data sets. The latter may be considered defined by five facets or dimensions: experimental units (persons), settings, time of measurement, treatment conditions and measurement variables.

Because cohort effects represent interindividual differences (across time, persons and settings), conceptualizing the cohort variable as a dimension of generalizability is a viable alternative to relegating it to a source of error or nondevelopmental variance. Obviously, identifying cohort as a facet of generalizability does not elevate it to the status of a full-fledged theoretical variable. But accepting it as another dimension of generalizability for the identification of interindividual differences in ontogenetic development implies that ensuring external validity will require the use of cohort-sequential research designs. Otherwise, researchers will be misled by the inadequate information supplied by simple longitudinal and cross-sectional work, which, in the case of simple longitudinal designs, does not provide coverage of this generalization facet or, in the case of cross-sectional work, confounds age with cohort differences.

An important added perspective in the treatment of cohort as a dimension of generalization is the distinction between quantitative (more or less) and qualitative (different process) generalization. It is now established that there is a great deal of cohort variation in level of behavior including the timing of onset or the range of variability (both inter- and intraindividual). However, as persuasively argued by McCall (1977), such variation does not immediately suggest developmental differences in process, mechanisms or functioning. In the research presented here, this distinction between quantitative and qualitative change is equally relevant. We find widespread differences in level but little evidence for cohort differences in structural patterns, whether in the domain of ability or personality.

Cohort as Theoretical or Process Variable

A third orientation views the cohort variable as a potentially major ingredient of developmental theory. Cohort is regarded as either an indicator of one or more hypothetical constructs which have not yet been fully delineated or as a construct which contains information about mechanisms of behavioral development.

Such an approach explicitly recognizes that cohort can be linked to a system of antecedent, process and consequent events to meet particular aims of description, explanation and modification of developmental change. Viewing cohort as a viable theoretical variable requires attention to the several dimensions involved in cohort explication: for example, the delineation of the form and nature of cohort change judged to be developmental or establishing the need for such concepts as stages or transitions in representing cohort change. Regarding cohort as a theoretical or process variable parallels Wohlwill's (1970, 1973) approach to the use of the age variable in developmental research and theory. Like age, cohort becomes part of the developmental variable indexing a cohort function.

In our view, treating cohort as a theoretical process variable is less likely for psychology than it is, for example, for sociology (Riley, et al., 1972; Riley, 1979) where the process of cultural change and the salience of historical dimensions are major foci for theory and research. However, there have been efforts in psychology as well to view cohort as a theoretical variable (Riegel, 1976). Baltes and his colleagues, for example, have identified three sources of influence on life-span development (Baltes, Reese & Lipsitt, 1980). One of them, labeled history-graded influences, derives from cohort-related research findings.

Sequential Strategies

In addition to these observations on theoretical context, a few remarks on design and analysis of cohort-related effects (via sequential strategies) are offered. The development and analysis of sequential strategies has a short but vigorous history. Our own view has been presented extensively (Baltes, et al., 1978, 1979; Schaie & Baltes, 1975). Our general position differs from others in that we (1) prefer to treat age and cohort as descriptive parameters, (2) argue that for developmental psychologists two components -- age and cohort -- are usually sufficient, thereby excluding time of measurement as an important parameter, and (3) hold that the notorious confounding of age, time of measurement and cohort effects is largely a conceptual artifact. We will briefly elaborate on these points.

Sequential strategies involve the orchestration of cross-sectional and longitudinal examinations of the ontogeny of successive generations. They provide a basis for estimating the relative significance of sequential ontogenetic (individual) and generational (historical) change components. Authors seem generally to agree that application of cross-sectional and longitudinal sequences as proposed by Baltes (1968; see also Schaie & Baltes, 1975) will result in the data points necessary for examining the

relationships between age, cohort and time of measurement. Cross-sectional sequences provide for the use of independent observations; longitudinal sequences are based on repeated observations across age. Schaie's (1965) three sequential strategies (cohort-, time- and cross-sequential) are not useful for identifying data collection schemes. Their focus is on data analysis.

Considerable disagreement and controversy has arisen concerning the proper forms of analysis and interpretation of sequentially gathered data. One area involves mathematical issues and questions of model identification and the other centers on aspects of substantive interpretation of the developmental components (age, cohort, time of measurement) potentially involved. The disagreements center largely on Schaie's proposition to distinguish for explanation purposes between three two-component models: cohort-sequential, time-sequential and cross-sequential. Schaie argued that, in addition to the descriptive value of his model, the three methods made developmental explanation possible because the three developmental parameters (age, cohort, time of measurement) could be linked to distinct sources of developmental change. Age effects could be tied to maturational antecedents, cohort effects to genetic and/or environmental antecedents and time of measurement effects to cultural factors. Schaie's three parameters then are assumed to be meaningful theoretical (explanatory) constructs. Whether or not this is useful is highly debatable. It is certainly not a necessary set of conditions.

In developmental psychology, it was exactly Schaie's concern for both descriptive data collection and explanatory data analysis that was the source of dissatisfaction for other researchers (e.g., Adam, 1978; Baltes, 1968; Buss, 1973). Baltes (1968; see also Price, 1976; Goldstein, 1979) rejected Schaie's categorical emphasis on treating age, cohort and time of measurement as meaningful theoretical entities. He proposed that Schaie's model be viewed as a purely descriptive one, that the cross-sectional and longitudinal sequences be designated the vehicles for gathering data to study developmental processes within the general framework and that for purposes of psychological research the age-cohort matrix be the preferred one. Furthermore, Baltes did not specify, as did Schaie, a particular form of data analysis and interpretation of effects. The primary rationale was that the three parameters (age, cohort, time of measurement) did not lend themselves to either a single theoretical interpretation or an unequivocal statistical estimate of their magnitude. Schaie and Baltes (1975) subsequently agreed that the distinction between the use of sequential strategies as data collection methods for the descriptive identification of development change versus their use as tools for explanatory data analysis helps one to understand the issue and disagreements involved. For example, it is not always understood

that Schaie's methods, contrary to Baltes' use of cross-sectional and longitudinal sequences, are not primarily methods of data collection. For Schaie, they are methods of data analysis and interpretation.

In case one decides, as did Schaie (1965), to elevate all three design parameters (age, cohort, time of measurement) to the level of meaningful theoretical constructs, as may be true for certain questions (usually not the kind developmental psychologists are interested in), there remains the problem area mentioned above, namely that of model identification and data analysis. Because the values of any two of the three design components define the third one (e.g., age equals time of measurement minus birth year), the effects associated with these classifications are inherently confounded. Researchers in other disciplines (e.g., Fienberg & Mason, 1979; Jackson, 1975; Mason, Mason, Winsborough & Poole, 1973; Price, 1976) have approached the identification and estimation problems with alternative schemes of analysis, but it appears that the statistical problem of unconfounding age, cohort and time of measurement effects remains to be solved in a completely general manner.

Because of widespread concern with statistical solutions of the age-cohort-time of measurement problem, we repeat our own views on the matter. With Glenn (1976), our position is that the statistical approach to the problem is inherently unsatisfactory, and except for rare exceptions unnecessary at least for the developmental psychologist. Thus, we have argued (Baltes, et al., 1979) that future use of sequential strategies in developmental psychology (but not necessarily in other disciplines such as sociology) will not primarily involve solving the confounded effects question by alternative schemes of data analysis. On the contrary, the central task is the formulation of a more fruitful conception of the meaning and utility of the variance components associated with an age by cohort matrix. A recent statement by Goldstein (1979) is also consistent with this viewpoint.

Why do we favor the use of age and cohort as the only design parameters for substantive and methodological reasons (see also Buss, 1973; Goldstein, 1979; Price, 1976) in addition to statistical ones? The developmental psychologist's task is to study intraindividual change and interindividual differences in intraindividual change. These two kinds of manifestations of developmental functions (Wohlwill, 1973) are most directly represented in an age by cohort data classification. Cohort differences represent one kind of interindividual difference in the nature of ontogenetic change (Baltes, Cornelius & Nesselroade, 1979). Therefore, in the typical case of developmental-psychological research, the identification problem disappears. The confounding issue can be regarded as a creation of researchers who, without sufficient justification

from a developmental psychology perspective, insist on three design parameters where two are usually sufficient for the task of identifying cohort-related differences in intraindividual change. In selected cases, however, it may be useful to acknowledge that all three design components should be given explicit attention. These are the special circumstances, however, that in our view need conceptual justification; not the reverse. It is inappropriate to always categorically assume that the statistical confound exists irrespective of theoretically considerations.

One may ask what we propose to do with the variance associated with time of measurement or how we would analyze data matrices which do not lend themselves easily to an age-cohort analysis. As to time of measurement, we recommend that the investigator examine a posteriori whether certain characteristics of age-cohort functions coincide with particular time periods intervening between observations. If that should be so, further search into conditions associated with these time periods is warranted. However, even in the case of time of measurement-specific effect patterns (the sociologist's "period" effects), we need to be careful not to infer that the antecedents are located in the time of measurement period studied. It is not necessarily so. Because of lagged and cumulative causation (Nesselroade & Blates, 1979), the appearance of a developmental phenomenon (its time of measurement) does not necessarily coincide with its causal origin.

As to analysis of sequential data, it is often true that simply applying an age-cohort model (requiring observations at the same age levels for all cohorts) would not exhaust all the information available in one's data matrix; for example, there might be observations on certain cohorts not consistent with the major age/cohort scheme employed. This occurs when one does not have a large number of occasions or when certain cohorts were observed only at particular age levels. In those instances, employment of data-analytic schemes that do not follow a complete age/cohort matrix is justified. In fact, this is true for the study to be reported here where age by time of measurement arrangements are used in order to maximize information available. These data analyses of incomplete age-cohort matrices, however, are approximations of the ideal. Moreover, in no instance do we assume that the analysis is aimed at identifying distinct sources of "explanatory" variances (genetic, maturational, environmental) as would be true for Schaie's (1965) approach.

A Sequential Study of Adolescent Personality

Objective

At the time the present study was designed, the examination of cohort effects in psychological variables had been focused largely on adult and elderly subjects, although some investigations had involved

the study of cognitive development in children (Baltes & Reinert, 1969). Moreover, previous work had centered mainly on measures of intelligence rather than personality. As a consequence, one of the central questions was whether research in other age groups (e.g., adolescence) and with noncognitive variables (e.g., personality) would yield similar evidence for the role of cohort effects in developmental psychology and the usefulness of sequential strategies.

We believe that the need to look for cohort (historical) components of change as well as age-related (individual) ones in conducting developmental research on adolescent personality is quite evident. Personality development during adolescence is assumed by most sociological and social psychological theorists to be particularly influenced by cultural change phenomena. Adolescents, in turn, are viewed as having a prominent role in shaping cultural change. Because in 1969 our interests were largely methodological and exploratory, we did not formulate a specific set of theory-guided hypotheses about relationships between cultural change and personality. Rather, our focus was on representatively mapping personality development of adolescents (age 13-18) occurring during a given historical period (1970-72).

If we had a theoretical conception, in addition to representative sampling of the personality domain, it was a set of pre-conceived notions about the general role of environmental conditions. We expected that those behavior classes which are largely determined by environmental and/or experiential conditions would exhibit the largest cohort variation. Personality variables (such as anxiety, achievement, ego strength, etc.) are among the classes of behavior that are generally assumed to be determined substantially by distinct learning histories and situational factors (e.g., Mischel, 1968; Vandenberg, 1966). In fact, in the few time-lag studies available at the time that were aimed at comparing related behavior systems such as attitudes and interests in different cohorts of adolescents (e.g., Broderick & Fowler, 1961; Greenstein, 1964; Harris, 1959; Jones, 1960), significant generational (cohort) change in adolescents had been found consistently. These findings provided some suggestive evidence of the susceptibility of personality variables to changes in societal conditions during the adolescent period of the life span.

Method and Procedures

Research Design
Design variables

Age, sex and cohort membership were varied following the strategy of sequential data collection (Baltes, 1968; Schaie, 1965). Longitudinal

sequences were used as the data collection strategy and resulted in the observations summarized in Table 1. Specifically, four short-term longitudinal studies, each involving three times of measurement, evenly spaced over a period of two years, were conducted. Note that, because only a two-year period was available, the four longitudinal studies did not cover the same age ranges. This is the reason why subsequent data analyses did not use an age by cohort scheme but employed time by cohort data arrays as an approximation to that arrangement.

Insert Table 1 about here

To provide a more powerful experimental design (Baltes, 1968; Baltes, Reese & Nesselroade, 1977), two sets of control groups were incorporated in the design. One was included so that testing effects could be assessed. It enabled us to estimate whether or not apparent changes (e.g., from 1970 to 1972) in the longitudinal group were confounded with effects of repeated testing. The control groups consisted of a new random sample, stratified by age and sex, drawn and tested at the third occasion of measurement (1972). The second set of controls was used to evaluate selective dropout effects. Those subjects who, although measured initially in 1970, did not complete the 1970-72 study, comprised the second control group. Comparison of this group's 1970 data with the 1970 data of the subjects who remained in the study reveals the extent to which selective mortality, with respect to our measurement variables and as reflected in the first-occasion data, has jeopardized the external validity (Baltes, Reese & Nesselroade, 1977; Campbell & Stanley, 1963) of the results.

Measurement System

Measuring instruments were chosen to represent the universe of personality variables from a structured measurement perspective. A small set of cognitive variables was included to mark the domain of human abilities. Table 2 provides an overview of the personality and ability measures used.

Insert Table 2 about here

Lacking a set of instruments corresponding to a particular developmental theory of adolescent personality, we decided to use measures from two batteries claimed by their authors to map personality structure in a

fairly comprehensive and psychometrically sound manner. Selected were Cattell and Cattell's (1969) High School Personality Questionnaire (HSPQ--Form A) and the Personality Research Form (PRF--Form E) of Jackson (1968). The two inventories differ sufficiently in both their underlying theory and development that together they should provide a fairly thorough mapping of the sphere of personality.

The HSPQ -- the adolescent version of the 16PF (Cattell, Eber & Tatsuoka, 1970) -- represents most distinctly the idea of convergence between a theory of personality structure and a corresponding set of measurement scales. Fourteen psychological concepts are measured by the HSPQ. Eight are regarded as being affected by age-development in middle childhood and adlescence, although the available studies are cross-sectional in nature and have failed to disentangle age from genera- tion effects (e.g., Sealy & Cattell, 1966). Jackson's (1968) PRF was developed using a multivariate convergent and discriminant validation approach to cover a broad spectrum of the behavioral universe. The PRF focuses essentially on the framework of Murray, et al. (1938), but with additional refinement of concept definition. As noted in Table 2, the subtests of the Primary Mental Abilities (PMA) battery (Thurstone & Thurstone, 1962) were used to represent the abilities domain.

Sample

Subjects were drawn from 32 junior and senior public high schools in three West Virginia counties located in the northwest of the state: Harrison, Wetzel and Wood. The base population included some 20,000 students. Stratified by grade, sex and homeroom unit, the sample was drawn at random from school rosters. The ratio of Caucasian to Negro population is about 95:5.

Summary information concerning the longitudinal, dropout and retest control groups is presented in Table 3. Approximately 2,000 students (from cohorts 1954-1957) were given a brief description of the study by letter and asked to participate. Of those 2,000 students, a total of 1828 participated in the personality analyses phase of the project and 1809 in the ability analyses phase. The initial volunteering rate for the 1954-1957 cohorts was approximately 91%; further, approximately 63% of the core longitudinal samples participated in all three times of measurement (1970, 1971, 1972).

Insert Table 3 about here

Testing

The tests were administered in classrooms during regular school periods. Groups ranged in size from about 30 to 90. Total testing time of four hours was divided into one morning (Primary Mental Abilities, High School Personality Questionnaire) and one afternoon (Personality Research Form) session. In a few cases, the two sessions had to be distributed over two days. The 1970, 1971 and 1972 occasions of measure- ment each extended over a period of approximately four months (late fall until early spring). January 1 was taken as mean testing date for all occasions.

Analyses and Results

The analyses presented here are illustrative of the relevance of cohort effects in developmental psychology and the role of sequential strategies in their identification. Analysis of cohort effects, like that of age differences, can concentrate on two main goals (Baltes & Nesselroade, 1973; Nunnally, 1973; Schaie, 1973). One is the assessment of difference in structure. The second is the analysis of age- or cohort-related differences in level.

For the most part, cohort-related research has focused on level. In fact, McCall (1977) lamented that this focus on quantitative rather than structural (qualitative) comparison has resulted in an overestimation of the role of cohort effects in psychological development. In the follow- ing, both structure and level are examined in relation to the major design varaibles. Analysis of structure is achieved by comparative factor- analytic work. Analyses of quantitative changes in level involves the use of ANOVA designs.

Structural Analyses

A very salient component of our investigation of these data is the examination of qualitative or structural change. Two kinds of analyses were conducted: (1) factor analytic comparisons of structures of measures; and (2) computation of stability (test-retest) coefficients.

There are two primary reasons why examinations of qualitative change or difference are important. First, stage sequences or other forms of ordered, qualitative change are often regarded as fundamental pillars of developmental theory (Riegel & Rosenwald, 1975). Second, the validity of assessments of quantitative change rests on the underlying stability

of one's measurement system (Baltes & Nesselroade, 1973). Without some evidence of a structurally stable framework, measurements taken at different times or developmental levels do not provide a convincing basis for comparisons followed by inferences about the nature of change. Nor do comparisons of measurements taken on different groups offer a sound basis for substantive interpretations of apparent between-group differences. In describing change over time, stability coefficients (test-retest correlations) add an additional dimension to one's interpretation of structural characteristics of data. Three stability coefficients were computed for each of the measures (10 personality and six ability dimensions, to be described later); two one-year stabilities (1970-71, 1971-72) and one two-year stability (1970-72).

How likely is it that we would find changes in structure in the present study? Obviously, one's underlying assumptions about the nature of development and developmental change play a significant role in the choice of measurement instruments. Our chosen measurement systems were quantitatively oriented and, thus, they are tilted towards structural invariance. But choosing such a framework does not foreclose completely the possibility of finding manifest qualitative change in the data. Nor of course does finding a good fit between data and quantitative change necessarily rule out the possibility that a satisfactory representation of development using a qualitatively oriented measurement framework might be achieved. Nevertheless, we acknowledge that the measurement framework chosen is not one that maximizes the probability of finding structural or qualitative differences.

Ability Dimensions

The age- and cohort-related structural characteristics of the ability measures were examined in two analyses (Fitzgerald, Nesselroade & Baltes, 1973; Hays, 1977). The central question is whether there are systematic changes in the patterning of abilities as a function of either age or cohort.

Fitzgerald, et al. (1973), working within the age-differentiation tradition (Reinert, 1970), examined the extent to which the factor structure of abilities as defined by the Primary Mental Abilities (PMA) test (Thurstone & Thurstone, 1963) provided to be invariant across independent samples of students chosen from three different age/grade levels (grades 7-8, 9-10, 11-12). The major strategy was one of cross-sectional age comparisons. All three groups yielded good simple structure factor patterns which corresponded closely to the pattern of adult intelligence underlying the construction of the PMA test. The outcome indicated that a developmentally robust organization of abilities was "in

place" prior to early adolescence. Hays (1977), in a similar vein, looked at the longitudinally obtained ability measurements. His longitudinal analysis outcomes supported the cross-sectional ones of Fitzgerald, et al., in showing no evidence of systematic difference in ability structures by cohort or sex grouping.

Thus, neither cross-sectionally nor longitudinally is there strong evidence of structural change in the pattern of abilities examined. On the contrary, the pattern of abilities obtained exhibited considerable invariance. This conclusion was further supported by evidence on longitudinal stability coefficients for the individual tests. For the most part, stabilities were high with a slight trend for highest stabilities to be found in the older age groups.

Personality Dimensions

Friel and Nesselroade (1976) examined the extent to which the HSPQ personality scales reflected the structure underlying the personality system embodied in the series of age-graded questionnaires stemming from the work of Cattell. These analyses, conducted separately on male and female data, provided evidence of stable structure in line with those expected a priori from knowledge of the measurement system. Furthermore, as to the PRF, Kafer (1977) did not find evidence of any dramatic structural differences or changes across data sets. Neither of these studies involved the systematic examination of structural changes but such work is currently in progress.

Note also that the stabilities of the personality measures used in subsequent analyses showed relatively consistent patterns. One year (1970-71, 1971-72) and two year (1970-72) stability coefficients for the ten personality dimensions were commputed separately by cohort and sex. The total range of stabilities was quite large (+.16 - +.81), although the average magnitude (.57) was fairly high for personality dimensions. The general pattern was one of a slight increment in stability with increasing age and decreasing stability as the time interval increases (1970-71, 1971-72 vs. 1970-72). In nine out of ten cases (except for Social-Emotional Anxiety), again for the total sample, the two-year stabilities were lower than both one-year stabilities. The increasing stability with age generally reflects a systematic ontogeny towards stable, trait-like rather than labile, state like-characteristics. For example, stability coefficients for the oldest cohort (1954) were nearly comparable to those for the ability-trait data.

Two additional findings of general interest emerged from the stability analysis. In nine out of ten cases (except for Factor VI -- Social-Emotional Anxiety), the 1971-72 stability coefficients were higher than

the corresponding ones for 1970-71. This could be due to average age differences, stabilization as a result of retesting and perhaps also a reflection of selective sampling. Higher longitudinal stabilities in 1971-72 than 1970-71 could be interpreted, however, also as indicating potential changes in the consistency of environmental influence patterns which mediate stability (see Baltes & Nesselroade, 1973, for a discussion of a developmental view of changes in stability indicators). Furthermore, there was a sex difference in stability. In pooled data, females showed higher stability (in 24 out of 30 cases) than did males.

Thus, in general terms, the analyses conducted thus far rather clearly point to a robustness of structure and of intraindividual differences during the period and for the cohort/age groups studied. This robustness of structure and stability of individual differences is more pronounced for abilities than personality. However, in absolute terms, the structural invariance and stability were fairly high for personality characteristics as well.

Quantitative Analyses

Data Reduction

An important preparatory effort was to evolve a set of measures which allowed a reduction in number of the 34 personality variables, but about which much was known concerning their structural characteristics before an attempt was made to examine ontogenetic and generational components of quantitative change. As to the six PMA measures, no further effort at data reduction was made. Ability tests were maintained intact for the analyses. As to personality, the 34 measures were reduced to 10 "second order" factors.

Specifically, for personality variables, the 14 HSPQ scales and the 20 PRF scales were first factor analyzed separately using the responses from 1877 subjects (cohorts 1951-57) for whom complete personality data were available at Time 1 (1970). Seven common factors were extracted from HSPQ scale intercorrelations and eight from PRF scale intercorrelations. These two sets of factors were then independently rotated to simple structure solutions. The two sets of factors (HSPQ and PRF) were then correlated with each other. A full account of the factor analytic procedure used and the outcomes is presented elsewhere (Nesselroade & Baltes, 1975).

The HSPQ and PRF exhibited both factors common to the two inventories and factors that were specific to one inventory. Substantial convergence was found in the case of four between-instrument pairs of factors. Each pair of the four HSPQ-PRF convergent factors was combined

into one dimension by adding appropriate pairs of estimated factor scores. In addition, each instrument had unique factors (HSPQ: Anxiety, Social Anxiety, Verbal Intelligence; PRF: Independence, Aggression, Achievement). This strategy of data reduction provided a parsimonious, developmentally robust assessment framework and organized data neatly into PRF-HSPQ common and HSPQ and PRF specific personality dimensions. The resulting personality dimensions are identified in Table 4.

Insert Table 4 about here

Control Analyses

Two kinds of analyses were conducted to seek information about the internal and external validity of the study. Because these control analyses revealed significant effects which should be taken into account in analyzing for developmental manifestations in level of ontogenetic versus cohort change, the outcomes from the control analyses will be presented first.

The first control analysis dealt with selective dropout. By means of a two (dropout vs. retestees) by four (cohort) by two (sex) ANOVA design we determined whether or not the core longitudinal sample and the dropout sample differed on any of the dependent variables at the first occasion of measurement (1970). The second control analysis was focused on retest effects. By means of a two (retestees vs. control) by four (cohort) by two (sex) analysis of variance of each personality and ability dimension, we checked for testing effects, using the third occasion (1972) as the check point.

The control analyses supplement each other because the retest control design is not ideal. This is so because the retest control group (tested once at the third occasion) is not fully comparable to the longitudinal sample (tested in 1970, 1971, 1972) in terms of volunteering behavior. Longitudinal subjects were expected to participate three times in the testing while controls were "in a holding pattern" until asked to participate in testing at the third occasion (1972). Thus, longitudinal subjects are likely to exhibit more dropout than the retest group. Therefore, the longitudinal data are potentially contaminated by testing and dropout effects, whereas the retest control groups' data should be less affected by dropout effects. In Table 5, a summary of findings from the retest and dropout control analyses is presented.

Insert Table 5 about here

As Table 5 shows, the outcomes are very different for personality and ability measures. Consider first the retest effects analyses. On all six ability measures, the longitudinal groups outperform their retest control counterparts. The two significant cohort by group interactions indicate that for Letter Series and Spatial Relations the effects are differential across cohorts. These findings indicate a substantial problem with retest effects in longitudinal observations involving ability tests. By contrast, only two of ten measures derived from the personality inventories showed significant retest effects. The longitudinal group was more extroverted and scored higher on the HSPQ verbal intelligence measure than did the controls.

Analyses for dropout effects were equally clearcut in outcome. For each of the ability measures, with the exception of Verbal Meaning, the core longitudinal sample was biased in a positive direction. Only one statistically significant effect was reflected in the personality measures. Dropouts scored higher than longitudinal subjects on the PRF Independence dimension.

The control analyses outcomes have several implications for the evaluation of the longitudinal subjects' data. First, the positive bias of the longitudinal subjects on ability dimensions indicates the need for restricting generalizations from sample to parent population. Second, retest control analyses indicated that apparent longitudinal changes in abilities are confounded with age (or time)-correlated, positive testing effects. Third, the dropout data suggest that the retest analysis itself is not fully internally valid since it involved comparing a positively biased longitudinal sample to a control group less affected by experimental mortality. Fourth, and perhaps most important for the present study, effects of testing and dropout are less substantial by far for personality than for ability measures.

Clearly, it would be desirable to correct these disparities, but the nature of the control effects is such that only approximate adjustments can be made. The nature of the adjustments made for the ability measures is illustrated in Figure 1. The procedure used to correct for biases was to deduct from the observed 1972 retest effect the amount due to dropout indicated by the 1970 dropout analysis. Because the retest effects were consistently positive and the longitudinal sample was biased positively when compared with the retest control group on each ability measure, this amounted to a subtraction in each case. (See Baltes, Reese & Nesselroade, 1977; Labouvie, Bartsch, Nesselroade & Baltes, 1974 for additional discussion of the rationale underlying the corrections.)

Insert Figure 1 about here

In sum, then, confounds reflected in dropout and retest control analyses were minor for the personality measures, but substantial for the ability measures used. It is apparent that the interpretation of simple longitudinal gradients as age change (internal validity) is not valid with ability measures given such dramatic retest effects. Moreover, again primarily for ability measures, the sample remaining in our longitudinal study is positively biased, suggesting a restriction in external validity.

Analyses for Quantitative Ontogenetic (Age) Versus Cohort Change

It was mentioned earlier that the analysis used in the 1974 SRCD monograph did not follow the ideal of an age by cohort arrangement. At the time, we judged the cohort by time matrix used to yield a satisfactory approximation to that ideal. Specifically, in Schaie's (1965) terminology, a cross-sequential method of data analysis was adopted for the investigation of the core longitudinal data. This arrangement varies cohort and time of measurement. It is shown in Table 6. Note again, however, that when using Schaie's term "cross-sequential" we do not imply its use as an explanatory model.

The total pool of subjects participating in the 1970, 1971 and 1972 data collection -- the sequential longitudinal group -- was divided into four levels of birth cohort (1954, 1955, 1956, 1957). A four (cohort) by two (sex) by three (time of measurement) analysis of variance of each personality and ability dimension was conducted to examine quantitative aspects of developmental change. Subsequently, the results are graphed to illustrate cohort-related variation of age functions. It would have been possible to employ alternative strategies of data analysis which would keep the measures of interest intact rather than further combine them in composite scores of very limited construct validity. ANOVA is one strategy which preserves the specificity of individual measures.

Insert Table 6 about here

As mentioned above, the 1974 SRCD monograph focused on time/cohort matrices in order to capture in one analysis as large a segment of the data matrix as possible. However, it was recognized that these time and cohort effects could be represented graphically as age functions for distinct cohorts, i.e., by our preferred age-cohort scheme. Thus, we used the overall analysis by time/cohort primarily to ascertain that significant patterns are present. Note, for example, that the time effect includes age variance. This is so because time variation (1970, 1971, 1972) is correlated with an increase in average age (12-15 vs. 13-16 vs. 14-17). Because our general position is that effects contained in any combination

of age, cohort and time of measurement designs should be construed to index age changes and cohort differences in age changes, we felt justified in subsequently charting the obtained effects as longitudinal age changes for distinct cohorts. This procedure, then, results in the following. First, the data matrices are analyzed via cohort-time arrangements; subsequently the data are charted as age changes for distinct cohorts.

Two additional observations are relevant. First, note that in the present cohort by time analysis, age effects are contained in both levels of cohort and time of measurement. This is the case because the oldest cohorts (by birthyear) are also older in age on the average (1954:17, 1955:16, 1956:15, 1957:14). For time of measurement the situation regarding average age is 1970:14 yr.,6 mo.; 1971:15 yr., 6 mo.; 1972:16 yr., 6mo. Thus, in the present scheme both cohort and time of measurement effects can indicate age differences. However, the age variation in cohort is larger (three years) than in time of measurement (two years). Therefore, if age would dominate the change pattern, cohort effects should dominate the outcome. But if time of measurement effects are more salient, historical, cultural change effects of the 1970-72 period should be more dominant.

Second, a major feature of the outcomes was that sex differences were found to be pervasive. This is an interesting finding, but not one of primary significance for the project since sex interacts only in three cases with either cohort or time. Sex effects were highly significant for all personality dimensions except Superego Strength/Impulse Control. Significant sex by time interactions were found for Tough-Mindedness/ Autonomy and Anxiety and a significant sex by cohort interaction was obtained for Intelligence. No significant triple interactions (cohort by sex by time) were found.

Ability Measures

Relative to the outcomes of the analyses of personality measures, those of ability measures are less interesting in the present context and we refer the reader to the 1974 SRCD monograph (Nesselroade & Baltes, 1974) for extensive information. In all analyses of variance on the sequential-longitudinal data, significant main effects of cohort and time of measurement were found. Sex differences were significant on three of the Primary Mental Abilities (NF, LS, SR). Only three of the 24 possible interactions (one sex by time, two cohort by time) reached significance. When the data are plotted against age, it is clear that birth cohort accounts for little variance; most of the variance (except for retest effects) is age-related and similar in trajectory for the four cohorts.

To illustrate more clearly the nature of the analyses performed and to present a more complicated situation, more detailed information on the analysis of Spatial Relations, a 30 item subtest of the PMA, is presented in Figure 2. Significant main effects were found for cohort, sex and time of measurement. The cohort by time of measurement interaction also was significant as shown. The sex main effect reflected higher scores for males ($\bar{x} = 33.7$) than females ($\bar{x} = 28.5$). The time effect indicated higher performance from 1970 ($\bar{x} = 23.8$) to 1971 ($\bar{x} = 31.7$) to 1972 ($\bar{x} = 37.3$). The significant difference between retest controls and longitudinal subjects amounted to 11.8 raw score points. A significant group by cohort interaction was found also. Means for the longitudinal group were 34.9, 38.0, 38.2 and 38.7 for cohorts 1957-1954, respectively. Corresponding means for the retest controls were 21.3, 27.0, 25.4 and 29.9. The retest effects are not reduced a great deal by dropout adjustments (from 11.8 to 9.5).

Thus, in this instance involving Spatial Relations, most of the apparent cohort-specific, longitudinal age change seems due to retest effects and not to ontogenetic change. In analyses of other ability measures, the retest effects are not as strong and clearer age-change patterns are obtained (Nesselroade & Baltes, 1974).

Insert Figure 2 about here

Personality Measures

In general, for personality measures, time effects were more prominent than cohort effects. Because of the relative degree of variation of age versus cohort in the present design (as described before), this result lends support to the interpretation that historical-cultural change effects covering the 1970-72 period are rather substantial.

The seven factors for which the main effects of time were found are Extroversion/Ascendance, Superego Strength/Impulse Control, Anxiety, Socio-Emotional Anxiety, Intelligence, Independence and Achievement. In contrast, significant cohort effects which, as stated before, in this design are primarily indicative of both prior-to-1970 cohort and age differences, were found for only two of the ten personality dimensions (Toughmindedness/Autonomy and Intelligence). Cohort and time interacted significantly, however slightly, in the cases of Superego Strength/Impulse Control and Achievement, but in general these interactions were few in number.

Thus, the analyses of personality dimensions suggests that, after sex effects, the 1970-1972 time dimension most systematically accounts for subgroup differences. Cohort effects are less frequent. Because cohort variation includes an average of three-year age difference, we believe that chronological age per se does not appear to be as powerful a variable as one might have expected. In other words, the nature of age change is different for distinct cohorts associated with the 1970-72 historical time period.

For purposes of illustration, three of the ten outcomes will be presented in greater detail. Each of the dimensions will be presented by name and a few representative adjectives given to remind the reader of the general nature of the factor. The data are represented as factor scores scaled to a mean of zero. The figures contain the cohort-specific longitudinal gradients (horizontal comparisons) determined over three occasions of measurement for each of the four cohorts. They also display separately for each occasion of measurement, cross-sectional age differences (vertical comparisons). The figures in Nesselroade and Baltes (1974) contain information, if necessitated by outcomes of control analyses, about the impact of testing effects. For the examples presented here, because of a lack of testing effects, no such information is given.

HSPQ-PRF Superego Strength/Impulse Control

Descriptors for this dimension include serious, conscientious, nonaggressive, inhibited. The means exhibited statistically significant time of measurement and cohort by time interaction effects. The means are plotted in Figure 3. Both males and females in the three youngest cohorts (1955-57) became less conscientious and controlled from 1970 to 1971 to 1972. The significant cohort by time interaction indicates that the systematic time-related decline in Superego does not apply to the oldest cohort (1954) which, statistically, exhibits no longitudinal change at all.

Insert Figure 3 about here

Furthermore, to obtain an indication of the magnitude of difference one would have obtained if various cross-sectional studies had been conducted, contrast the four cohort-specific longitudinal gradients (horizontal comparisons) with the findings of the three cross-sectional (vertical comparisons) representations applying to the 1970, 1971 and 1972 measurements. In 1970, the younger the adolescent, the more conscientious and controlled he or she tended to be. By 1971, however, the

older adolescents were the ones who scored highest on Superego Strength/ Impulse Control. Similar discrepancies between various cross-sectional outcomes and between cross-sectional and longitudinal gradients apply to many of the personality dimensions studied.

PRF Achievement

Striving and persistent are two of the chief descriptors for the Achievement dimension. It manifested significant main effects of sex and of time of measurement and a significant cohort by time interaction. Females scored lower on Achievement than males. Means, reflecting the cohort by time interaction, are presented in Figure 4. The general pattern is one of lowered Achievement from 1970 to 1972 for the two youngest cohorts and one of relative "steady state" for the two oldest cohorts. Younger adolescents (from 1970 to 1972) apparently were exposed to socialization conditions that led to reductions in achievement scores. This longitudinal ontogenetic finding does not appear to apply to the older adolescents during this same time period.

Insert Figure 4 about here

Again, dramatic discrepancies between various cross-sectional and between longitudinal and cross-sectional gradients can be seen in Figure 4. Achievement scores for 14-year-old adolescents in 1970 contrast sharply with those for 14-year-olds in 1971 and 1972. Mean scores for 14-year-olds across the three time periods range nearly as widely as the means for all age levels during the same time period. In many cases, different age adolescents (e.g., 14 vs. 15) observed at the same point in time, say 1972, are more similar to each other than they are to same age adolescents measured at any of the other two times of observation. Because sex does not interact with other design components, Achievement seems to be another personality dimension on which sex differences are established prior to adolescence.

PRF Independence

The terms autonomous, non-recognition-seeking and secure convey the flavor of this personality dimension. Significant main effects of sex and of time of measurement were found and are illustrated in Figure 5. Males scored higher than females ($\bar{x}$ = .36, -.41, respectively) but more dramatic is the increase in independence shown by all cohorts (independent of their age) over the two-year period. The outcome, again, is one of marked differences between cross-sectional studies and different levels for same age groups from different cohorts. As with Achievement, the lack of interaction with sex indicates the prior emergence and maintenance of the sex difference found.

Insert Figure 5 about here

Discussion and Conclusions

One aspect of discussion will be centered on a number of issues pertaining to the examination of age- and cohort-related change components. We wish to note at the outset that the generalizability of the present findings is subject to limits set by the nature of the subject population, the measurement battery, the procedures followed, etc. Given our findings concerning the impact of cultural change on development, it is clear that one of the dimensions along which generalization is most limited is that of time periods. One period -- 1970 to 1972 -- is a questionable representative of time periods in general, especially in a rapidly changing culture.

A second, more abstract focus of discussion involves implications of the results of this study for research design and theory building in developmental psychology. A growing number of theoretical papers have reflected critical evaluations of mechanistic, personological and epoch-centered world views that seem to characterize much developmental research (e.g., Baltes & Schaie, 1976; Bengston & Black, 1973; Elder, 1975; Hartup & Lempers, 1973; Overton & Reese, 1973; Reese & Overton, 1970; Riegel, 1976). Developmental models which have emphasized an active organism in a passive environment or a passive organism in an active environment are seriously challenged by dialectic notions in which development and change of the organism and environment are viewed as concurrent, mutually influencing phenomena (Lerner & Ryff, 1978). Strong empirical support for such views is not yet available but the writers believe that outcomes such as the present ones testify to the salience of interactive relation-ships between individual and historical change components in developmental change.

Implications for Conceptions of Adolescent Personality

Age Versus Cohort/Time Effects in Structure

The dominant outcome of the structural analyses was one of invari-ance. This is a particualrly noteworthy point for developmentalists because structural invariance highlights an island of stability in what is sometimes regarded to be a sea of change. McCall (1977), it was noted earlier, recognized the importance of examining various historical (secular, generational, cohort, etc.) influences on development but

argued that they could tend to be manifested primarily in changes or differences in score level rather than in changes or differences in the pattern or structure of interrelationships among variables. McCall's expectation is true for the present situation. While there is much cohort-related variation regarding age development (as a function of the 1970-72 historical moment) in terms of quantitative indices of rate and directionality, such variation does not extend to the structure of abilities and personality. The distinction between level and structure is key not only from the standpoint of the theorist who would like to account for developmental change but from a practical standpoint as well.

The outcome of relative structural invariance is further supported by longitudinal analysis of stability. By and large, stability of inter-individual differences is high. However, in addition, our data suggest that stability is somewhat less interesting as an indicator of reliability or prediction than as a representative of developmental variation. Based on the data presented here, a given measure shows not one characteristic level of stability but rather a range of stabilities according to age, sex, cohort membership and time of observation. For example, the fact that females tend to show higher stability than males warrants attempts to identify underlying gene-environment mechanisms. Furthermore, given that 1971-72 produced higher age-related stability than 1970-71 suggests, for example, that patterns of socialization during the later period main-tained individual differences of trait-like personality attributes more than those of the earlier period. Similarly, the age-related differences in personality stability (increases with age) support the idea that trait-like behavior patterns emerge from more labile classes of behaviors (Baltes & Nesselroade, 1973).

Age Versus Cohort/Time Effects in Level

The outcome with regard to level differed, in the case of person-ality, from that found in structure comparisons. There are major cohort-related differences in personality level. While for abilities, the main finding is one of testing and dropout effects, for personality functioning the 1970-1972 time epoch, except for existing sex differ-ences, accounts for a major portion of the variation in adolescents' functioning. This outcome provides something of a challenge to notions of orderly, stage-like, sequentially unfolding patterns of behavior development during adolescence.

The consistent discrepancies which we found between cross-sectional and longitudinal data in personality, for example, do not support the invariant developmental trends portrayed in many biologically and per-sonologically oriented models of adolescent development. Rather, average quantitative standing and ontogenetic trends seem more dependent on the historical time (cultural moment) to which the adolescents have been exposed. In other words, for adolescents, level of personality functioning

is less defined by sequences of chronological age than by sequences of historical change. Whether this phenomenon can be generalized to other populations, measuring instruments, research situations, etc., remains for future examination.

Given that considerable influence on the developmental process is exogenous to the organism and that these exogenous determinants are not necessarily unidirectional and cumulative, the need for developmentalists to enlarge their conceptual framework beyond traditional disciplinary boundaries is apparent. Social psychological perspectives on adolescence (e.g., Keniston, 1970; Lerner & Spanier, 1950; Muuss, 1962) have asserted the significance of ecological and societal influences on adolescents for some time but it is only with the application of newer designs such as cross-sectional and longitudinal sequences that insight into society-ontogeny relationships can be obtained. Similar concerns about the acceptance of unidirectional and personological socialization models as opposed to interactive relationships have been expressed in other substantive areas as well, for example, in work on parent-child relationships (e.g., Bronfenbrenner, 1977). Similarly, the present authors (Baltes et al., 1979) have contributed to a heuristic, multidisciplinary model of human development which elucidates the joint role of three systems of influences: age-graded, history-graded and nonnormative events. Findings on cohort-related variability were a major influence in the development of this multicausal and interdisciplinary formulation.

Substantive Interpretation of Time/Cohort Effects

We view the primary usefulness of sequential strategies to be descriptive, to help establish the phenomenon of cohort variation in age changes. If one were interested in a research program aimed at causal-analytic interpretation of cohort variation, alternative and supplemental designs would need to be employed. Such work would use, for example, measures of historical contexts, strategies of cohort simulation (Elder, 1979; Labouvie, 1978), or techniques of causal modeling (Nesselroade & Baltes, 1979).

For illustrative purposes, however, we shall briefly speculate about the substantive origin of the cohort/time effects observed in the present study, but we do it in full recognition that these attempts at interpretation are highly tentative. The pattern of obtained variations in age development and our general view of the 1970-1972 socio-cultural context in the United States guide our speculation.

It appears to us, although in principle such a possibility exists, that the variation in age trends for the four cohorts involved is difficult to explain by genetic or biological processes. It seems more

parsimonious to assume that the pattern of differential cohort changes with age originates in experiential processes common to most subjects. Obviously, such common experiences could be either a reflection of participation in the study (instrumentation or testing) or of shared cultural influences. Due to the lack of testing effects in the case of personality measures, we tend to conclude that for personality dimensions (the focus of this report), the observed time effects signify notable cultural change arising from socio-cultural conditions of the 1970-1972 time period. Thus, the observed 1970-1972 decrement in adolescent Superego Strength and Achievement and the correlated increase in Independence are interpreted to reflect historical socio-contextual development.

Our information does not permit the pinpointing of specific correlates for change in socio-cultural contingencies from 1970 to 1972. The focus on aggressive behavior displayed by wide segments of the American society in conjunction with youth activism and the Vietnam War, the tendency of youth to occupy itself with ethical, moral and political issues rather than cognitive achievement, reports by various public polls of gradual decline in respect for and confidence in public and educational leadership, however, are elements of the cultural context prevailing at the time. A history-related decline in Superego Strength and Achievement and a counterpart increase in Independence would parallel such a "syndrome" of socio-cultural context.

It should be noted again that this effort at interpreting the reasons for the changes we observed is not compelling from a research design perspective. The available evidence is not only correlational. In addition, we could not even locate a careful, systematic description of the type of socio-cultural change that characterized the living context of our subjects.

Adolescent Development and Sex Differences

The pattern of outcomes on sex differences not only supports the generally recognized pervasiveness of sex differences but adds some new perspectives to the issue as well. Significant differences between males and females were found on seven of the ten personality and on five of the six ability dimensions, but the number of significant interactions of sex with either time or cohort were few.

There is some evidence that sex differences shown at age 12 did tend to become more, rather than less, pronounced during this period. But the dominant finding is that the observed sex differences, though quite pervasive, were established prior to age 12 for the cohorts we studied and do not seem to be affected by either age or the historical context of the 1970-72 period. The 1970-72 time period involved little redirecting or accentuating of sex differences in the traits we studied.

This outcome, together with the fact that female stability coefficients tended to be higher than those for males, does not suggest the gradual dissolution of sex role boundaries over the historical period examined that one finds occasionally portrayed in the nonscientific literature on activism and feminism. Yet, note that the adolescents studied tended to live in more rural areas rather than those that might be characterized as centers of adolescent activism.

Generalizability

It has been mentioned already that it is important to acknowledge that the findings reported here are specific not only to our samples but also to the historical period. Because of the danger of elevating our findings to a "general" phenomenon and because such a conclusion would run counter to our basic position, we risk redundancy on this score and offer some additional observations.

Recognizing generalization over times of measurement to be an important concern is significant because it would be detrimental to our research if it were concluded that historical periods always exercise the particular influence pattern reported here. Only further research using sequential strategies could examine the long term nature of historical change and its relationships to adolescent personality development. There might be historical periods when the effects obtained are negligible, others in which the effects are in the opposite direction of those presented here. In fact, we would expect a diversity of patterns to occur periodically. How one would conceptualize, as a psychologist, the nature of cohort-related historical variation is also a function of one's treatment of the cohort variable. As mentioned in the opening sections of this chapter, several options exist (error, temporary disturbance, quantitative generalization, theoretical process).

Findings which offer an interesting comparison to ours were reported by Schaie (1979). Age-related increases in some dimensions of intellectual behavior were found to be specific to certain historical periods for older age subjects. For example, during the period 1956-1963 apparent effects of socio-cultural change occur at a slower rate than during the period 1963-1971 as reflected in unspeeded performance measures. The differences are particularly marked for individuals who are in their 40s and 50s at those times of measurement. Diverse findings on cohort variation and conceptualization thereof, then, are part of the game.

Implications for Developmental Research Design

Simple Versus Sequential Designs

From the standpoint of both age versus cohort/time effects and the outcome of control analyses it seems clear that neither simple cross-sectional nor the often praised simple longitudinal design is

satisfactory for descriptive developmental research. The pervasiveness of cohort/time differences is once again supported by the results reported here. The data clearly imply that cohort or historical differences obtrude into the measurement particularly of self-reported personality. Adolescence is implicated as a period in which cohort/time differences must be taken into account if one wants to obtain accurate descriptions of developmental change. Substantial cohort/time effects were observed even though the temporal definition of cohort was reduced to the short span of one year.

In this context, it needs to be emphasized that the alternative to cross-sectional design is not simple longitudinal studies. Simple longitudinal designs do not provide information of broad generality involving the age development of multiple cohorts. Different cohorts and historical times can show marked differences in the direction and rate of age-related changes whether these cohort-related differences extend to analyses of structural (or qualitative) differences in ontogeny is an open question at this point. Furthermore, as to internal and external validity, the rather spectacular retest and dropout effects in ability measures argues against the general usefulness of simple longitudinal designs.

Developmental researchers must become more aware of these critical methodological issues. If a descriptive data base from which theoretical advance can proceed is to be generated, appropriate remedial steps in research procedures must be taken. Sequential strategies (cross-sectional and longitudinal sequences) with appropriate controls can be argued to provide a much more appropriate general design format than either simple cross-sectional or longitudinal methods.

As a concluding observation, we stress that future research on behavioral and developmental variation among different cohort groupings will require investigators to choose methods and analysis procedures to match their conception of the role of cohort as a variable. In general, we do not believe that cohort variation is always of theoretical interest to developmental psychology. On the contrary, as was mentioned before, it is quite possible that cohort variation can be taken as unwanted error or disturbance. However, even if such an atheoretical view of cohort effects is accepted, it is necessary to conduct cohort-sequential research in order to capture the "unwanted" and to be able to focus on cohort invariant developmental functions. Thus, the use of sequential strategies is a must. To the dialectically oriented developmental psychologist, they provide the "real stuff"; to the personologically oriented researcher, they tell what not to study.

References

Adam, J. Sequential strategies and the separation of age, cohort, and time-of-measurement contributions to developmental data. Psychological Bulletin, 1978, 85, 1309-1316.

Baer, D.M. An age-irrelevant concept of development. Merrill-Palmer Quarterly, 1970, 16, 238-246.

Baltes, P.B. Longitudinal and cross-sectional sequences in the study of age and generation effects. Human Development, 1968, 11, 145-171.

Baltes, P.B., Cornelius, S.W. & Nesselroade, J.R. Cohort effects in behavioral development: Theoretical and methodological perspectives. In W.A. Collins (Ed.), Minnesota symposia on child psychology (Vol II). Hillsdale, New Jersey: Erlbaum, 1978.

Baltes, P.B., Cornelius, S.W. & Nesselroade, J.R. Cohort effects in developmental psychology: In J.R. Nesselroade & P.B. Baltes (Eds.), Longitudinal research in the study of behavior and development. New York: Academic Press, 1979.

Baltes, P.B. & Goulet, L.R. Explanation of developmental variables by manipulation and simulation of age differences in behavior. Human Development, 1971, 14, 149-170.

Baltes, P.B. & Nesselroade, J.R. The developmental analysis of individual differences on multiple measures. In J.R. Nesselroade & H.W. Reese (Eds.), Life-span developmental psychology: Methodological issues. New York: Academic Press, 1973.

Baltes, P.B., Reese, H.W. & Lipsitt, L.P. Life-span developmental psychology. In M.R. Rosenzweig & L.W. Porter (Eds.), Annual review of psychology, Vol. 31, Palo Alto, CA: Annual Reviews, Inc., 1980.

Baltes, P.B., Reese, H.W. & Nesselroade, J.R. Life-span developmental psychology: Introduction to research methods. Monterey, California: Brooks/Cole, 1977.

Baltes, P.B. & Reinert, G. Cohort effects in cognitive development of children as revealed by cross-sectional sequences. Developmental Psychology, 1969, 1, 169-177.

Baltes, P.B. & Schaie, K.W. On the plasticity of intelligence in adulthood and old age: Where Horn and Donaldson fail. American Psychologist, 1976, 31, 720-725.

Bell, R.Q. & Hertz, T.W. Toward more comparability and generalizability of developmental research. Child Development, 1976, 47, 6-13.

Bengtson, V.L. & Black, K.D. Intergenerational relations and continuities in socialization. In P.B. Baltes & K.W. Schaie (Eds.), Life-span developmental psychology: Personality and socialization. New York: Academic, 1973.

Broderick, C.B. & Fowler, S.E. New patterns of relationships between the sexes among preadolescents. Marriage and Family Living, 1961, 23, 27-30.

Bronfenbrenner, U. Toward an experimental ecology of human development. American Psychologist, 1977, 32, 513-531.

Buss, A.R. An extension of developmental models that separate ontogenetic changes and cohort differences. Psychological Bulletin, 1973, 80, 466-479.

Campbell, D.T. & Stanley, J.C. Experimental and quasi-experimental designs for research on teaching. In N.L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

Cattell, R.B. & Cattell, M.D.L. Handbook for the Jr.-Sr. High School Personality Questionnaire "HSPQ". Champaign, Ill.: Institute for Personality and Ability Testing, 1969.

Cattell, R.B., Eber, H.W. & Tatsuoka, M.M. Handbook for the sixteen personality factor questionnaire (16PF). Champaign, Ill.: Institute for Personality and Ability Testing, 1970.

Damon, A. Discrepancies between findings of longitudinal and cross-sectional studies in adult life: Physique and physiology. Human Development, 1965, 8, 16-22.

Elder, G.H., Jr. Age differentiation and the life course. Annual Review of Sociology, 1975, 1, 165-190.

Elder, G.H., Jr. Historical change in life patterns and personality. In P.B. Baltes & O.G. Brim, Jr. (Eds.), Life-span development and behavior (Vol. 2). New York: Academic Press, 1979.

Emmerich, W. Personality development and concepts of structure. Child Development, 1968, 39, 671-690.

Fienberg, S.E. & Mason, W.M. Identification and estimation of age-period-cohort models in the analysis of discrete archival data. In K. Schuessler (Ed.), Sociological methodology 1979, San Francisco, California: Jossey-Bass, 1979.

Fitzgerald, J.M., Nesselroade, J.R. & Baltes, P.B. Emergence of adult intellectual structure: Prior to or during adolescence? Developmental Psychology, 1973, 9, 114-119.

Friel, J.C. & Nesselroade, J.R. Test of the replicability of Cattell's HSPQ factor structure using item parcels. Multivariate Experimental Clinical Research, 1976, 2, 23-32.

Glenn, N.D. Cohort analysts' futile quest: Statistical attempts to separate age, period, and cohort effects. American Sociological Review, 1976, 41, 900-904.

Goldstein, H. Age, period and cohort effects -- A confounded confusion. BIAS, 1979, 6, 1-24.

Greenstein, F.I. New light on changing American values: A forgotten body of survey data. Social Forces, 1964, 42, 441-450.

Harris, D.B. Sex differences in the life problems and interests of adolescents, 1935 and 1957. Child Development, 1959, 30, 453-459.

Hartup, W.W. & Lempers, J. A problem in life-span development: The interactional analysis of family attachments. In P.B. Baltes and K.W. Schaie (Eds.), Life-span developmental psychology: Personality and socialization. New York: Academic Press, 1973.

Hays, W.S. Stability of factor structure of intellectual abilities in adolescents: A maximum-likelihood factor analytic investigation of the age-differentiation hypothesis. Unpublished Masters thesis, College of Human Development, The Pennsylvania State University, 1977.

Jackson, D.J. A reformulation of Schaie's model of developmental change. Paper presented at the meeting of the Gerontological Society, Louisville, Kentucky, October, 1975.

Jackson, D.N. Personality Research Form. New York: Goshen, 1968.

Jones, M.C. A comparison of the attitudes and interests of ninth grade students over two decades. Journal of Educational Research, 1960, 51, 175-186.

Kafer, R.A. Examination of the structure of adolescent personality development by confirmatory factor analysis. Unpublished Masters thesis, College of Human Development, The Pennsylvania State University, 1977.

Kagan, J. & Moss, H.A. Birth to maturity: A study in psychological development. New York: Wiley, 1962.

Keniston, K. Postadolescence (youth) and historical change. In J. Zubin & A.M. Freedman (Eds.), The psychopathology of adolescence. New York: Grune & Stratton, 1970.

Keniston, K. Psychological development and historical change. Journal of Interdisciplinary History, 1971, 2, 330-345.

Kerlinger, F.N. Foundations of behavioral research. New York: Holt, Rinehart and Winston, 1964.

Kuhlen, R.G. Age and intelligence: The significance of cultural change in longitudinal vs. cross-sectional findings. Vita Humana, 1963, 6, 113-114.

Labouvie, E.W. Experimental sequential strategies for the exploration of ontogenetic and socio-historical changes. Human Development, 1978, 21, 161-169.

Labouvie, E.W., Bartsch, T.W., Nesselroade, J.R. & Baltes, P.B. On the internal and external validity of simple longitudinal designs. Child Development, 1974, 45, 282-290.

Lerner, R.M. & Ryff, C.A. The implementation of the life-span view of human development: Sample case of attachment. In P.B. Baltes (Eds.), Life-span development and behavior (Vol. 1). New York: Academic Press, 1978.

Lerner, R.M. & Spanier, G.B. Adolescent development: A life-span perspective. New York: McGraw-Hill, 1980.

Maccoby, E.E. (Ed.), The development of sex differences. Stanford: Stanford University Press, 1966.

Mason, K.O., Mason, W.M., Winsborough, H.H. & Poole, W.K. Some methodological issues in cohort analysis of archival data. American Sociological Review, 1973, 38, 242-258.

McCall, R.B. Challenges to a science of developmental psychology. Child Development, 1977, 48, 333-344.

Mischel, W. Personality and assessment. New York: Wiley, 1968.

Murray, H.A., et al. Explorations in personality. Cambridge, Mass.: Oxford University Press, 1938.

Muuss, R.E. Theories of adolescence. New York: Random House, 1962.

Nesselroade, J.R. & Baltes, P.B. Adolescent personality development and historical change: 1970-1972. Monographs of the Society for Research in Child Development, 1974, 39 (1, Serial No. 154).

Nesselroade, J.R. & Baltes, P.B. Higher factor convergence and divergence of two distinct personality systems: Cattell's HSPQ and Jackson's PRF. Multivariate Behavioral Research, 1975, 10, 387-408.

Nesselroade, J.R. & Baltes, P.B. (Eds.), Longitudinal research in the study of behavior and development. New York: Academic Press, 1979.

Neugarten, B.L. & Datan, N. Sociological perspectives on the life cycle. In P.B. Baltes & K.W. Schaie (Eds.), Life-span developmental psychology: Personality and socialization. New York: Academic, 1973, in press.

Nunnally, J.C. Research strategies and measurement methods for investigating human development. In J.R. Nesselroade & H.W. Reese (Eds.), Life-span developmental psychology: Methodological issues. New York: Academic Press, 1973.

Overton, W.F. & Reese, H.W. Models of development: Methodological implications. In J.R. Nesselroade & H.W. Reese (Eds.), Life-span developmental psychology: Methodological issues. New York: Academic Press, 1973.

Price, D.O. A respecification of variables in cohort analysis. Unpublished manuscript, University of Texas, 1976.

Quetelet, A. A treatise on man and the development of his faculties. Edinburg: William and Robert Chambers, 1842.

Reese, H.W. & Overton, W.F. Models of development and theories of development. In L.R. Goulet & P.B. Blates (Eds.), Life-span developmental psychology: Research and theory. New York: Academic Press, 1970.

Reinert, G. Comparative factor analytic studies of intelligence throughout the human life-span. In L.R. Goulet & P.B. Baltes (Eds.), Life-span developmental psychology: Research and theory. New York: Academic Press, 1970.

Riegel, K.F. The influence of economic and political ideology upon the development of developmental psychology. Psychological Bulletin, 1972, 78, 129-141.

Riegel, K.F. The dialectics of human development. American Psychologist, 1976, 31, 689-700.

Riegel, K.F. & Rosenwald, G.C. (Eds.), Structure and function: Developmental and historical aspects. New York: Wiley, 1975.

Riley, M.W. (Ed.), Aging from birth to death. Boulder, Colorado: Westview Press, 1979.

Riley, M.W., Johnson, W. & Foner, A. (Eds.), Aging and society (Vol. 3) A sociology of age stratification. New York: Russel Sage, 1972.

Ryder, N.B. The Cohort as a concept in the study of social changes. American Sociological Review, 1965, 30, 843-861.

Schaie, K.W. A general model for the study of developmental problems. Psychological Bulletin, 1965, 64, 92-107.

Schaie, K.W. A reinterpretation of age-related changes in cognitive structure and functioning. In L.R. Goulet & P.B. Baltes (Eds.), Life-span developmental psychology: Research and theory. New York: Academic Press, 1970.

Schaie, K.W. Methodological problems in descriptive developmental research on adulthood and aging. In J.R. Nesselroade & H.W. Reese (Eds.), Life-span developmental psychology: Methodological issues. New York: Academic, 1973.

Schaie, K.W. The primary mental abilities in adulthood: An exploration in the development of psychometric intelligence. In P.B. Baltes & O.G. Brim, Jr. (Eds.), Life-span development and behavior. Vol. 2. New York: Academic Press, 1979.

Schaie, K.W. & Baltes, P.B. On sequential strategies in developmental research: Description or explanation? Human Development, 1975, 18, 384-390.

Sealy, A.P. & Cattell, R.B. Adolescent personality trends in primary factors measured on the 16PF and the HSPQ Questionnaires through ages 11-23. British Journal of Social and Clinical Psychology, 1966, 5, 172-184.

Thurstone, L.L. & Thurstone, T.G. SRA Primary Mental Abilities. Chicago: Science Research Associates, 1962.

Vandenberg, S.G. Contributions of twin research to psychology. Psychological Bulletin, 1966, 66, 327-352.

Wohlwill, J.F. Methodology and research strategy in the study of developmental change. In L.R. Goulet & P.B. Baltes (Eds.), Life-span developmental psychology: Research and theory. New York: Academic Press, 1970.

Wohlwill, J.F. The study of behavioral development. New York: Academic Press, 1973.

Table 1

Longitudinal Sequences Design* and Data Collection Scheme

		Time of Measurement		
		1970	1971	1972
Cohort	Sex			
1954	M F	16	17	18
1955	M F	15	16	17
1956	M F	14	15	16
1957	M F	13	14	15

* Entries are ages of subjects at particular times of measurement. January 1 of each year was the mean testing time ($\pm$ 2 months). Samples of randomly selected retest control subjects were drawn from cohorts 1954-57 and tested for the first time in 1972. The core longitudinal sample and dropouts were contrasted on 1970 scores to estimate dropout effects.

Table 2

Variables Measured in Primary Mental Abilities Test, High School Personality

Questionnaire and Personality Research Form (after Nesselroade & Baltes, 1974)

Instrument	Variables	
I. Primary Mental Abilities (Thurstone & Thurstone, 1962)	1. Verbal Meaning	2. Word Grouping (Reasoning)
	2. Number Facility	5. Number Series (Reasoning)
	3. Letter Series (Reasoning)	6. Spatial Relations
II. High School Personality Questionnaire (Cattell & Cattell, 1969)	1. Sizothymia (A)	8. Parmia (H)
	2. Intelligence (B)	9. Premsia (I)
	3. Ego Strength (C)	10. Coasthenia (J)
	4. Excitability (D)	11. Guilt Proneness (O)
	5. Dominance (E)	12 Self-sufficiency (Q_2)
	6. Surgency (F)	13. Self-sentiment (Q_3)
	7. Superego (G)	14. Ergic Tension (Q_4)

Table 2 (Continued)

Instrument	Variables	
	1. Abasement	11. Exhibition
	2. Achievement	12. Harmavoidance
	3. Affiliation	13. Impulsivity
III. Personality Research	4. Aggression	14. Nurturance
Form	5. Autonomy	15. Order
(Jackson, 1968)	6. Change	16. Play
	7. Cognitive Structure	17. Sentience
	8. Defendence	18. Social Recognition
	9. Dominance	19. Succorance
	10. Endurance	20. Understanding

Note -- Each of the HSPQ scales contains 10 items, whereas the PRF scales consist of 16 items each. The number of items included in the PMA scales varies: Verbal Meaning, $N = 60$; Space, $N = 30$; Number Facility, $N = 30$; and Reasoning, $N = 70$ (consisting of three subscales: Letter Series, $N = 20$; Word Grouping, $N = 30$; and Number Series, $N = 20$). Letters in parentheses following HSPQ variables are designations used by Cattell to identify personality dimensions.

Table 3

Sample Size for Sequential Longitudinal, Dropout, and Retest Control Groups Separately for

Personality and Ability Analyses (after Nesselroade & Baltes, 1974)

Cohort	Sequential-Longitudinal (1970-1972)		Dropout (1970-1972)		Retest Control (1972)		Total N
	Male	Female	Male	Female	Male	Female	
Personality Variables							
1957	102	119	35	37	70	86	449
1956	95	127	64	56	70	72	474
1955	101	123	87	67	50	66	483
1954	83	66	82	79	60	52	422
Ability Variables							
1957	99	118	44	37	68	84	450
1956	93	123	66	48	69	71	470

Table 3 (Continued)

Cohort	Sequential-Longitudinal (1970-1972)		Dropout (1970-1972)		Retest Control (1972)		Total N
	Male	Female	Male	Female	Male	Female	
			Ability Variables				
1955	99	118	89	69	48	52	475
1954	80	63	86	79	56	50	414
Total							
Personality	381	435	268	229	250	265	1828
Total Ability	371	422	285	233	241	257	1809

Table 4

Ten Personality Dimensions Obtained From Factoring

HSPQ and PRF Scales*

Factor		Saliently Loading Scales
I. Extraversion/Ascendance	HSPQ:	A+, F+, H+, Q_2-
	PRF:	Dominance (+), Exhibition (+), Sentience (+)
II. Superego Strength/Impulse Control	HSPQ:	F-, G+, Q_3+
	PRF:	Aggression (-), Cognitive Structure (+),
		Harmavoidance (+), Impulsivity (-),
		Order (+), Play (-)
III. Tough Mindedness/Autonomy	HSPQ:	A-, E+, I-, Q_2+
	PRF:	Nurturance (-), Sentience (-), Understanding (-)
IV. Independence/Avoidance of Social Contact	HSPQ:	J+, Q_2+
	PRF:	Affiliation (-), Change (-), Exhibitionism (-),
		Play (-), Understanding (-)
V. Anxiety	HSPQ:	C-, D+, Q_2-, Q_4+
VI. Social Anxiety	HSPQ:	C-, O+

127

Table 4 (Continued)

Factor		Saliently Loading Scales
VII. Intelligence	HSPQ:	B+
VIII. Independence	PRF:	Autonomy (+), Change (+), Harmavoidance (-), Social Recognition (-), Succorance (-), Understanding (+)
IX. Aggression	PRF:	Abasement (-), Aggression (+), Defendence (+), Understanding (+)
X. Achievement	PRF:	Achievement (+), Dominance (+), Endurance (+), Harmavoidance (-), Understanding (+)

* A complete account of the procedures leading to the development of the factors presented here and the factor loadings and intercorrelations is available elsewhere (Nesselroade & Baltes, 1975).

Table 5

Significant Outcomes of Retest and Dropout Control Analyses Separately by Personality and Ability Dimensions

Dimensions		Retest Control (1972)			Dropout Control (1970)		
		G	G X C	G X S	G	G X C	G X S
HSPQ-PRF	I	*					
	II						
	III						
	IV						
HSPQ	V						
	VI						
	VII	*					
PRF	VIII				*		
	IX						
	X						
PMA	VM	*					
	NF	*			*		
	LS	*	*		*		
	WG	*			*		
	NS	*			*		
	SR	*	*		*		

NOTE: 1% level of confidence was used for comparisons. C = cohort, S = sex, G = group; G always involves the longitudinal group and the retest control or the longitudinal group and the dropout control. (VM = Verbal Meaning, NF = Number Families, LS = Letter Series, WG = Word Group, NS = Number Series, SR = Spatial Relations)

Table 6

Cross-Sequential Analysis of Adolescent Development: Age, Cohort, And time
of Measurement (after Nesselroade & Baltes, 1974)

| Mean Cohort | Number of Subjects | | | | Mean Time of Measurement | | |
| | Personality | | Ability | | | | |
	M	F	M	F	January 1970	January 1971	January 1972
May 1957	102	119	99	118	12:6	13:6	14:6
May 1956	95	127	93	123	13:6	14:6	15:6
May 1955	101	123	99	118	14:6	15:6	16:6
May 1954	83	66	80	63	15:6	16:6	17:6

NOTE: Table entries on right are approximate ages (years:months) at the three times of measurement. Range of each cohort and age level is an interval of one year. Analyses of variance involved cohort by time by sex designs thus covarying chronological age in the sense of Schaie's (1965) cross-sequential model.

Figure Captions

Figure 1. Method of control analysis-based score adjustment. Apparent retest effect, d_1, is reduced by dropout effect, d_2, to estimate longitudinal gradient.

Figure 2. Longitudinal sequences for Spatial Relations ability. Significant main effects of time and cohort and cohort x time interaction were found. Magnitude of retest effects is indicated by arrows. (After Nesselroade & Baltes, 1974).

Figure 3. Longitudinal sequences for Superego Strength/Impulse Control. Significant main effect of time of measurement and cohort x time interaction were found. (After Nesselroade & Baltes, 1974).

Figure 4. Longitudinal sequences for Achievement showing main effect of time of measurement and cohort x time interaction. (After Nesselroade & Baltes, 1974).

Figure 5. Longitudinal sequences for Independence showing main effect of time of measurement. (After Nesselroade & Baltes, 1974).

Chapter 6

Structural Equation Models in Longitudinal Research

P. M. Bentler

University of California, Los Angeles

Structural Equation Models in Longitudinal Research

P.M. Bentler

University of California, Los Angeles

There is little question that structural equation models represent the "cutting edge" of methodology for dealing with those longitudinal research problems that can be represented as a system of linear influences among variables across time. These models are particularly important when they deal with latent constructs as well as with the measurement operations that relate the constructs to measured variables, for then they are able to disentangle theoretically meaningful influences of constructs on each other from the relatively uninteresting effects of random errors of measurements. As a consequence, theories can be tested with nonexperimental data using a relatively well-developed and statistical methodology. Unfortunately, this methodology is quite complicated, and the substantive researcher will have to study it because, unlike many other methods, "causal modeling" cannot be relegated to assistants or statisticians for implementation. The modest goal of this chapter is to provide an introduction to structural equation models. A more comprehensive overview is given in Bentler (1980), where a relatively complete bibliography of relevant research can also be found. This chapter, however, provides more examples to illustrate the modeling process. See also Kenny (1979).

Development and Role of Structural Equation Models

General linear latent variable causal models represent the convergence of research traditions in psychometrics, econometrics, and biometrics. The concepts of latent variables or unmeasured variables, and errors in variables, have had the longest tradition of development in psychometrics, in which they have come to be known as factor analysis and reliability theory. Similarly, the simultaneous directional influences of some variables on others have been studied intensively for decades in econometrics, primarily with manifest or observed variables under the label of simultaneous equation models. Finally, a specialized tradition in biometrics, associated primarily with Wright (1934), has dealt with simultaneous equation models in the context of representation and estimation schemes known as path analysis. These traditions remained relatively independent and unaware of each other until the 1960's, when sociological methodologists such as Blalock (1961), Boudon (1965), and Duncan (1966) demonstrated the value of combining the simplicity of path analytic representations with the rigor of specifying equations simultaneously. By the early 1970's, causal modeling was a major sociological research method (Blalock, 1971), and latent variable models were being studies in depth. Excellent perspectives on these developments are provided by the econometrician Goldberger (1971, 1972), who has become a major supporter and contributor to the field (see Goldberger & Duncan, 1973; Aigner & Goldberger, 1977). Psychologists were not major contributors to these developments. While Campbell and Stanley (1963) had begun to focus on problems of causal inference in nonexperimental data (e.g., with cross-lagged panel correlation), the main introduction of these ideas into psychology is due to Werts and Linn (1970). However, this work did not inspire extensive imitation. In my opinion, this occurred in large part because unifying mathematical and statistical principles, and simple procedures for their implementation, had not yet been developed for the methodology.

Many publications up to the early 1970's followed Simon (1954) in providing highly detailed analyses of artificial and small problems. However, such approaches were too limited in scope to provide a broad enough framework to solve more realistic problems. What was needed was a general framework, analogous to multiple factor analysis or analysis of variance, that could in principle deal with extremely complex models in a routine way. Such a framework was provided by the JKW model (Jöreskog, 1973a, 1977; Keesling, 1972; Wiley; 1973). The JKW model (widely known as LISREL) appears to have been conceived by several individuals, but the statistician-psychometrician Karl G. Jöreskog, in my opinion, deserves to be recognized as the major developer of hypothesis-testing methods for analysis of nonexperimental data, particularly via latent variable models. He provided the first practical computer implementation of the statistical approach to factor analysis (Jöreskog, 1967, 1969), and to a generalized version of Thurstone's second-order model (Jöreskog, 1970, 1973b). His LISREL computer implementation of the JKW model has become the standard of the field (Jöreskog & Sörbom, 1978). Jöreskog has achieved an important balance between statistical sophistication and concern for relevance to social science applications (Jöreskog, 1978, 1979).

Generalizations and simplifications of the JKW model exist today (Bentler & Weeks, 1979a). Although important psychometric and statistical issues in causal modeling remain to be solved, the field has progressed to the point where quite general causal structures can be dealt with on a routine basis without requiring investigators to study a whole host of seemingly unrelated techniques to deal with special situations. For example, general latent variable models can also deal with manifest variable causal models. Consequently, it is not surprising that there exists a growing consensus about the relevance of causal modeling to such areas as economics (Aigner & Goldberger, 1977), education (Anderson & Evans, 1974; Cooley, 1978), evaluation research (Bentler & Woodward, 1979), political science (Alker, 1969), and sociology (Bielby & Hauser, 1977). The reception psychology, nonetheless, remains slow: a special issue on methodology of the Journal of Consulting and Clinical Psychology (Aug. 1978) does not

even mention the topic. This is unfortunate, since latent variable causal modeling finally provides a basis for quantifying and operationalizing well-known concepts of construct validity and nomological networks as spelled out by Cronbach and Meehl (1955) and Torgerson (1958); (Bentler, 1978).

Major Structural Equation Concepts

Structural Equations and Path Diagrams

The basic building block of a causal model is the linear regression equation (linear by assumption). Such an equation specifies the hypothesized effects of certain variables (here called predictors) on another variable (here called criterion). To illustrate, consider the equation $Y=b_1X_1+b_2X_2+b_3X_3+e$. In such an equation, the intercept term has been dropped as irrelevant, and one considers the four variables Y, X_1, X_2, and X_3 as deviations from their means. (To be strictly accurate, one could add a subscript to the variables indicating the scores of a given entity or subject; there would be as many equations as entities. However, these equations are identical in form and are governed by the same parameters, so only one generic equation is needed). The parameters b_1, b_2, and b_3 represent the regression weights to be used in optimally predicting Y from the Xs, and e represents an error of prediction. The variable e is not actually measured; in the population, however, one would know the weights b_i, and hence e could be calculated exactly as the residual $(Y-b_1X_1-b_2X_2-b_3X_3)$. I shall call such variables unmeasured, but not latent. In this equation, there are four predictor variables X_1-X_3 and e_1, and Y is the criterion variable.

A path diagram for this equation is shown in Figure 1a. Squares are used to enclose variables that are measured (manifest variables); the unmeasured variable e has a circle inside the square. The predictors are shaded, and the criterion is light. Causal or directional influences of predictors on the criterion are indicated by unidirectional arrows; the strength of each effect is indicated by the weight for each arrow. The

diagram can be read as indicating that "$\underline{Y}$ equals $\underline{b}_1$ times $\underline{X}_1$, plus $\underline{b}_2$ times $\underline{X}_2$, plus $\underline{b}_3$ times $\underline{X}_3$, plus 1.0 times $\underline{e}$," thus completely summarizing the equation.

A regression equation in the context of a causal model is called a structural equation, and the parameters, structural parameters. Structural parameters presumably represent relatively invariant parameters of a causal process, and are considered to have more theoretical meaning than ordinary predictive regression weights. A problem with structural equations is that they do not adequately represent the parameters of a causal process. Implicit in each equation are parameters associated with the variances of the predictor variables (here, σ^2_1, σ^2_2, σ^2_3, σ^2_e) as well as their covariances (here, σ_{12}, σ_{13}, σ_{23}, since the residual $\underline{e}$ is forced to be independent of the $\underline{X}$s by construction). Hence, there are more parameters associated with a causal process than are represented in the structural equation. In the example, there are 10 parameters in the causal system, but only three of them are shown in the equation. Figure 1b presents a more complete representation of the model. Although such a representation is not standard, it mirrors the causal process more accurately. Covariances or correlations among the predictor variables are shown by two-headed arrows, and variances are marked inside the squares.

Typically, only the form of the model is known, and the parameter values need to be estimated from the data. Often it is desirable to test hypotheses about given parameters, e.g., that $\underline{b}_1$=0. If this null hypothesis cannot be rejected, the path diagram would need to be redrawn, and the arrow corresponding to $\underline{b}_1$ would be removed. Such a revision of the path diagram points out an important feature: missing paths are as important to accurate representation as existing paths. In Figure 1, the absence of a connection between $\underline{X}_1$ and $\underline{e}$ shows that these variables are neither causally dependent, nor correlated. If $\underline{X}_1$ and $\underline{X}_2$ were independent, the σ_{12} two-headed arrow would be missing.

Causal models typically have more than one equation. Figure 2a shows a confirmatory factor analysis model that proposes that there are two common factors ($\underline{X}_1$, $\underline{X}_2$), which may have a nonzero covariance (σ_{12}), as well as four independent error or unique factors. Latent variables (latent variables) are simply common or unique factors. The six latent variables are represented in circles, the four manifest variables in squares. If $\underline{Y}_1$ and $\underline{Y}_2$ represented mathematical tasks, while $\underline{Y}_3$ and $\underline{Y}_4$ represented verbal tasks, $\underline{X}_1$ and $\underline{X}_2$ might be hypothesized quantitative and verbal factors of intelligence. In such a model, the common factors generate the correlations among the manifest variables, and the manifest variables are indicators of the latent variables; note that the arrows go from latent variables to manifest variables. This diagram follows typical practice in not showing the variances of the predictors, which are parameters of the model. The predictors $\underline{X}$ and $\underline{e}$ are independent in each of the four equations, ($\underline{Y}_1=\underline{b}_1\underline{X}_1+\underline{e}_1$, $\underline{Y}_2=\underline{b}_2\underline{X}_1+\underline{e}_2$, $\underline{Y}_3=\underline{b}_3\underline{X}_2+\underline{e}_3$, $\underline{Y}_4=\underline{b}_4\underline{X}_2+\underline{e}_4$). Since there are no two-way arrows, the $\underline{e}$'s are independent across equations. In all latent variable models, one common factor regression weight must be fixed as known to identify each factor (see below), or else the factor variance must be fixed; thus, one can set b_1=b_3=1.0 and let σ^2_1 and σ^2_2 be free, unknown parameters, or else one can set σ^2_1= σ^2_2=1.0 and let $\underline{b}_1$-$\underline{b}_4$ be free parameters (in this case, σ_{12}=$\underline{r}_{12}$, a correlation). Hence, there are 9 parameters in the model (e.g., with factor variances fixed, there are 4 variances, 1 correlation, and 4 regression weights), one less than the model of Figure 1. An alternate theory for the manifest variables might be that there is only a single common factor (say, general intelligence) rather than two separate but correlated factors. Such a theory would be verified in current representation if $\underline{X}_1$ and $\underline{X}_2$ were perfectly correlated, i.e., if σ_{12}=r_{12}=1.0. Then the $\underline{X}$ part of the diagram could be redrawn to show a single common factor $\underline{X}$ with an arrow pointing to each $\underline{Y}_1$.

In the previous models, each variable was either a predictor or criterion. In Figure 2a, the four manifest variables are always criteria, never predictors. More complicated models allow variables to serve both

as predictors. More complicated models allow variables to serve both as predictors and criteria in different equations. An example is given in Figure 2b. The manifest variables are X_1, X_2, Y_1, and Y_2, while e_1 and e_2 are unmeasured variables. Variable Y_1 serves as a criterion, since it is predicted by X_1, X_2, and the unmeasured residual e_1. However, it also serves as a predictor of Y_2. I shall call those variables that never serve as criteria in any structural equation independent variables; the remaining variables are dependent variables. Then, there are exactly as many structured equations in a causal model as dependent variables. As many structured equations in a causal model as dependent variables. Independent variables are shaded in the figures, dependent variables remain light; thus there are two equations in the last-described model. Using this convention, one can describe the complexity of a causal model by the number of equations, the number of independent variables, the number of unknown regression weights, and the number of unknown, nonzero interrelations among the independent variables. (Note that the independent variables need not be mutually statistically independent or uncorrelated.)

The examples illustrate structural equations and path diagrams in the context of manifest variable path and patent variable factor models. In factor analytic models, the latent variables are independent variables only; more general latent variable models would have latent variables as dependent variables also. If the model in Figure 2b represented relations among theoretical constructs only, one would replace the squares by circles. In addition, one would need to provide manifest variables as indicators for each of the latent variables, using unidirectional arrows emanating from each latent variable to several manifest variables, with their residuals.

Model Specification and Research Design

A theory that is to be tested via causal modeling will have to be specified mathematically, i.e., translated into structural equations. Such a specification should assure that all the relevant constructs are

being considered simultaneously, and that their uni- and bi-directional interrelations are made explicit. Ideally, a competing theory should be specified at the same time, since the only clean test of competing theories can be made when one theory can be represented as part of another, i.e., their path diagrams are identical except that certain paths can be taken as known (usually, zero). Such nested comparisons can be assessed statistically. For example, in Figure 2a, two-factor and one-factor theories can be compared by evaluating whether $\sigma_{12} = r_{12} = 1.0$; a third two-factor theory might propose that $\sigma_{12} = 0$, i.e., that the factors are independent; yet another theory might propose that all regression weights b_1, ..., b_4 are equal.

The theoretical relations between constructs can be specified without anchoring the constructs in measurement operations. Such a specification must represent a reasonable translation of theory into equations or a diagram. If the scientific community cannot be convinced about the logic of the representation, there is no point in relating the constructs to indicators. If the theoretical representation is adequate, care must still be taken to provide adequate indicators of each construct. Each construct is considered to be a common factor latent variable that affects several manifest variables; each manifest variable, in turn is also affected by a residual latent variable (e's in Figure 2a) that is typically not an explicit part of a theory. Since the latent variables are in practice abstractions that presumably underlie manifest variables, a poor choice of manifest variables will create doubt as to whether a theory's constructs are in fact embedded in a model. Choosing the right number of indicators for each latent variable is something of an art: in principle, the more, the better; in practice, too many indicators make it difficult if not impossible to fit a model to data. While one might hypothesize that certain manifest variables represent only a given latent variable, manifest variables tend not to behave as well as expected. Since linearity of relations and continuity of variables are typical causal modeling assumptions, manifest variables should be chosen with these properties in mind.

Statistical tests also typically are based on independence of observations and normality of manifest variables; these conditions should be met if possible.

The structural equations associated with a causal model imply some very specific consequences for the moment structure of the data, specifically the variances and covariances of the manifest variables (if the manifest variables are standardized, covariances are correlations). Thus causation implies correlation -- but of a very specific form. If the hypothesized causal process is correct, only certain values will be observed for these variances and covariances (which is why structural equation models often are called covariance structure models). For example, in Figure 2a, if a model specifies that $\sigma_{12}=0$ and $\sigma_1^2 = \sigma_2^2$ are fixed at 1.0, then the covariance of Y_1 with Y_2 must equal the product $b_1 b_2$. This may be seen as follows. Assuming that Y_1 and Y_2 are standardized, this model proposes that X_1 generates the correlation between Y_1 and Y_2. If this model is true, partialing X_1 out of Y_1 and Y_2 should leave these manifest variables uncorrelated. But the partial correlation formula, rearranged, states that $r_{12.3} \{(1-r_{13}^2)(1-r_{23}^2)\}^{\frac{1}{2}} r_{12}-r_{13}r_{23}$ (where 1,2,3 are Y_1, Y_2,X_1). If the hypothesis $r_{12.3}=0$ is true, it follows $r_{12}=r_{13}r_{23}$. Then, with $r_{13}=b_1$ and $r_{23}=b_2$, the conclusion follows: Implications of this sort are drawn for every variance and covariance of the manifest variables, based on any assumed model. Any observed covariances other than the expected ones would be inconsistent with the proposed model, thus providing a basis for rejecting a model. However, even if a model is consistent with data, one cannot conclude that it mirrors the true causal process, since other models also might be able to reproduce the moment structure of the data.

Not all causal models can be tested. Models can only be tested if the parameters of the model can be uniquely specified or identified. In the example, it was possible to conclude that $r_{12}=r_{13}r_{23}$ only because the variance of X_1 was fixed at 1.0; it would not also be possible to treat this value as unknown, to be determined. The parameters of a model are identified if all parameters would have identical values in various causal

models that generate identical observed data. In simple models it is possible to evaluate the identification problem by finding transformations that take the manifest variable moments into parameter values, but this task is virtually impossible in large models and one typically uses certain heuristics instead (e.g., residual e variables in Figures 1 and 2 have 1.0 paths to other variables). A useful causal model must be overidentified, meaning, loosely speaking, it should have fewer parameters than data points (usually variances and covariances), because only then is the model potentially able to be rejected by data. If a model is just identified, meaning, loosely speaking, that there is a one-to-one transformation possible between parameters and data, the model is not scientifically interesting because it can never be rejected. The regression model of Figure 1, taken by itself, is such a model, since it can be fit to any data; only if certain parameters are taken as known (e.g., $b_1=.5$) would a potentially rejectable model exist. On the other hand, if the manifest variables were replaced by latent variable in Figure 1, one would have a model for regression with latent variables and several factor analytic indicators were chosen for each latent variable in Figure 1, one would have a model for regression with latent variables that could be tested. If the parameters of a model are under identified, meaning, loosely speaking, that they can take on many values rather than be determined uniquely, the model is not statistically testable and thus scientifically useless. Since certain parts of a model may be overidentified and other parts underidentified, identification is an issue for every equation in a model as well as for the model as a whole. The analysis of identification, while difficult, can provide insight into possible deficiencies of a theory (Bentler, 1978).

Model specification thus should be closely related to research design. Arbitrarily gathered data usually will not have the characteristics needed to provide an adequate test of a theory. It is necessary to assure that the conditions for data gathering are theoretically appropriate and statistically adequate. If theory specifies that certain influences occur only across time, with a certain causal lag, the time of measurement must reflect

this lag; if various populations are expected to differ in specific ways, it is necessary to use a multiple population modeling method; if an identification problem exists, use of additional manifest variables may cure the problem; and so on. Enough subjects must be tested to assure that conditions for evaluating theory are appropriate. Not only is the statistical theory applicable primarily to large samples, but one can guarantee the inability to reject a model by using a small enough sample. This is an inappropriate way to make a theory look good! Unfortunately, the researcher is to some extent in a double bind, since with extremely large samples almost any theory will be rejected by current procedures.

Model Testing

Once a model has been specified and the relevant data gathered in the context of an appropriate design, it is possible to compare the hypothesized model to date. The raw data are irrelevant to this process. Only the variances and covariances typically are utilized, though means are relevant in some context. The unknown parameters of the model are estimated so as to make the variances and covariances (possibly, means) that are reproduced from the model in some sense close to the observed data. Obviously, a good model would allow very close approximation to the data. If even the best choice of parameter values lead to a poor approximation to the data, the model can be rejected as a plausible representation of the causal process that generated the data.

In general it is impossible to estimate model parameters without a computer program because there is no available algebraic solution, and iterative approximations that refine a user-provided initial solution are utilized. The programs COFAMM and LISREL (Sörbom & Jöreskog, 1976; Jöreskog & Sörbom, 1978) use the maximum likelihood method of estimation, while the author's program additionally uses generalized least squares and least squares criteria; the latter has no statistical test associated with it. These programs may yield meaningless results even when the user sets up the problem correctly: an identification problem may preclude obtaining

a solution; the initial values may be so far off optimum that the program may not converge on a final solution; and some parameter estimates may be completely unreasonable (i.e., negative variances). An appropriate statistical solution will yield a chi-squared (χ^2) value to evaluate the goodness-of-fit of the model to data, and standard errors to reflect the sampling variability of each parameter estimate. The χ^2 statistic provides a test of the proposed model against the general alternative that the manifest variables are simply correlated to an arbitrary extent. If the χ^2 is large compared to degrees of freedom, one concludes that the model does not appropriately mirror the causal process that generated the data. If the statistic is small compared to degrees of freedom, one concludes that the model provides a plausible representation of the causal process. The standard error for each parameter estimate can be used to provide an indication of the importance of that parameter to the model as a whole. If the critical ratio formed by dividing the estimate by its standard error is large, the parameter is essential to the model; if it is small, the parameter is probably unnecessary to the model. These critical ratios have an approximate z distribution, so that the standard normal curve provides the index for deciding between "large" or "small." Parameter estimates can also be transformed into values that would be obtained if the common factor latent variables have unit variances, yielding a standardized solution.

In very large samples, the most trivial discrepancy between model and data will require rejection of a model by the χ^2 test. This test has drawbacks, since it is affected by the extent to which crucial assumptions such as multinormality and linearity are violated. Even if all assumptions are met, one might wish to take the significance level chosen to evaluate a model as a decreasing function of sample size (Leamer, 1978). Unfortunately, no standard methodology exists for such choices. The goodness-of-fit of a model to data should certainly be evaluated by methods besides the χ^2 statistic, for example, by examining residuals or evaluating a coefficient that does not depend on sample size (Tucker & Lewis, 1973).

In such evaluations, one frequently wants to show that a model provides a plausible representation of data. This is difficult to do with statistical hypothesis testing procedures, since it entails accepting the null hypothesis that the model provides a plausible representation of the data. Within such a framework, statistical power plays a paradoxical role.

Any competing model can be estimated and evaluated as stated above. However, if one can specify an alternative model that is a subset of the initial model, the difference in x^2 values between the two models is itself a x^2 statistic which can be used to test the importance of the parameters that differentiate the models. When there is no alternative model, and a model does not fit, one may wish to modify the model in a heuristic manner.

Model Modification and Reconfirmation

Parameters whose estimates are small compared to their standard errors can be eliminated from a model, and the resulting model reestimated. This process amounts to modifying the path diagram by removing paths. Paths can also be added by examining the residuals, i.e., the specific patterns of lack of fit of model to data. Certain derivatives can be examined (Sörbom, 1975). Parameters in a causal model are embedded in a matrix representation, and zero entries in various matrices correspond to missing paths in a diagram. If derivatives of the fit function are large with respect to these missing paths, it is possible that adding the associated parameters may improve the fit. Additional paths representing the correlations between errors can also be added. While procedures such as these usually produce a new model with better fit to data, it should be recognized that they can capitalize on change associations in the data. They will also not necessarily find an alternative model that might provide a far superior fit.

Cross-validation provides an appropriate way of establishing whether empirically based model modifications represent genuinely valuable information about a model. For example, a sample may be split in two halves, and one half used to develop a model and the second to provide a clean test of the developed model. It is possible to use tight, moderate, and loose replication strategies (G.J. Huba, J.A. Woodward, P. M. Bentler & J.A. Wingard, unpublished). In tight replication, one would attempt to fit the model to the second sample using the first sample's exact parameter estimates. In loose replication, the identical model and fitting procedures are used in both samples. In moderate replication, critical theoretical parameters (such as factor loadings) are held constant but otheres (such as error variances) can be estimated in the new sample. Research is required to differentiate these methods, but factor invariance theory (Bloxom, 1972) would favor the moderate strategy.

Examples of Structural Equation Models

Test of Theories of Orgasmic Responsiveness

Masters and Johnson (1966) reported that orgasm represents a single process that all women experience to a greater or lesser extent, regardless of the manner of stimulation. The contrasting idea that orgasm may not represent a uniform response has a long history that goes back at least to Freud; in recent years such writers as Singer and Singer (1972) and Fisher (1973) have espoused the idea. Bentler and Peeler (1979) hypothesized that women's responsiveness to coital and masturbatory stimulation are based on two distinct dimensions of subjective experience defined by coital and masturbatory responsiveness, and they contrasted such a model with the Masters and Johnson model of unidimensionality. Self-report ratings of subjective responses to orgasm were made by 281 female university undergraduates. Principal components analysis was used to reduce the initial ratings to two dimensions. Items marking each dimension were combined to

yield two scale scores for each dimension. There were thus four scale scores, and the covariance matrix of these scales was subjected to confirmatory factor analysis. A variety of model comparisons were made, of which the most relevant comparison is reported here.

The path diagram for this situation resembles that of Figure 2a with minor modification. Scales Y_1 and Y_2 were hypothesized to be indicators of masturbatory responsiveness (X_1). Scales Y_3 and Y_4 were hypothesized to be indicators of coital responsiveness (X_2). (In addition to these two latent variables, and third latent variable (X_3) underlying Y_1 and Y_3 was hypothesized; it is not drawn in the figure, nor relevant to the present discussion since it is a clearcut "method" factor, uncorrelated with X_1 and X_2, due to the response format of the rating scales; while it is present in the analyses, it has no bearing on the crucial model contrast.) The factors X_1 and X_2 were identified by setting their variances to unity. On the basis of the way the scales were constructed, and some previous model comparisons, the equalities $b_1=b_2$, $b_3=b_4$ and equality of all error variances were imposed. There were five parameters to be estimated (1 weight for each of 3 factors, 1 error variance, and the correlation σ_{12}). The four manifest variables have 10 variances and covariances, and thus there are 10-5=5 degrees of freedom (d.f.). The chi-square associated with this solution was 4.01, yielding an associated probability of .55. Thus, the two-dimensional theory of orgasmic responsiveness could not be rejected, and it represents a plausible model for the data. In this solution, masturbatory and coital factors were found to correlate .34. The Masters-Johnson position would predict that this correlation should not differ significantly from 1.0. Consequently, the model was respecified, with the only change being that $\sigma_{12}=1.0$ was imposed. There are thus 4 parameters to be estimated, and 6 d.f. This model was found to yield $X^2=1396.92$ with $p<.0001$. Thus, the 10 observed manifest variances and covariances would be extremely unlikely to be obtained in random sampling from a population in which the unidimensional theory was true. Stated differently, the unidimentional theory can be rejected. The difference in chi-squares between the two models

yielded 1396.92-4.01 - 1392.91, which provided a 1 d.f. test of the hypothesis that subjective masturbatory and coital responsiveness were perfectly correlated. This hypothesis could obviously be rejected. Note that the crucial model comparison involved a hypothetical correlation among latent variables that were not actually measured. An inference based on the correlation between observed scale scores would be biased, because errors of measurement would of necessity reduce any observed correlation below 1.0.

Bentler and Peeler (1979) also developed a manifest variable structural equation model that is consistent with the hypothesis that heterosexual and monosexual behaviors act as mediators between extraversion, neuroticism, and attitudes toward masturbation, on the one hand, and subjective coital and masturbatory responsiveness on the other hand. They discussed some therapeutic implications of the path model.

Comparison of Models of Attitude-Behavior Relations

Fishbein and Ajzen (1975; Ajzen & Fishbein, 1977) have advanced a major theoretical statement on the interrelations among attitudes, subjective norms, intentions, and behavior. Assuming certain research design considerations to hold, they proposed a model that can be represented as in Figure 2b. With the notation X_1 = attitude, X_2 = subjective norm, Y_1 = intention, and Y_2 = behavior, they proposed that attitudes and subjective norms influence future behavior only through the mediation of the intention to perform the behavior. Note that there are no direct influences of attitudes on behavior, nor of subjective norms on behavior. Such a model is theoretical, of course, and hence the manifest variables (squares) in the figure should be replaced by latent variables (circles); to make it operational, multiple indicators of each latent variable would be needed. Bentler and Speckart (1979) proposed some alternative models for these relations into which the Fishbein-Ajzen model can be embedded. In one model, a significant direct effect of attitude on future behavior (i.e., a

path from X_1 to Y_2) was hypothesized. In another model, past behavior was hypothesized to be a significant predictor of intentions to engage in a behavior as well as future behavior itself. Such a model is shown in Figure 3, which clearly differentiates between the theoretical constructs (circles) and manifest variables (squares). In this model, initial behavior (B_1), attitudes (A), and subjective norms (SN) are considered to be independent variables, allowed to correlate freely. Attitudes are shown to have a direct effect on future behavior (B_2), and initial behavior a direct effect on intensions (I) and future behavior; the Fishbein-Ajzen theory would predict that these three effects are not necessary to the model. In addition to the three shaded latent variables, errors of measurement and prediction are independent latent variables in the system. These 17 latent variables are indicated by arrows without a source.

Bentler and Speckart (1979) obtained questionnaire data about alcohol, marijuana, and hard drug use from 228 subjects on two occasions. These behavior domains were chosen to allow all model tests to be replicated three times. Each domain was represented by 15 variables, and each latent variable by three manifest variables in the manner shown in Figure 3. The three manifest variables were chosen to generate some variation in type or context of substance usage, so that the latent variable would represent the shared content of three manifest variables.

The simple Fishbein-Ajzen model (in latent variables, as in Fig. 2) was tested first and compared to a model with an additional path from attitudes to behavior; thus, B_1 and its indicators and consequences (as in Fig. 3) were excluded from these analyses. The Fishbein-Ajzen model could be rejected for two substances ($p<.05$), but it was marginally acceptable for one ($p=.05$). However the model with a direct attitude-behavior effect was acceptable in all cases ($p<.05$), and a X^2 difference test evaluating the improvement in fit due to the additional path showed that it provided a significant increment in fit of model to data ($p<.001$) in all cases. The complete model of Figure 3 showed an acceptable fit to the data in all

replications, so that it provides a plausible representation of attitude behavior relations. In this model the 15 manifest variables have 15(16)/2=120 variances and covariances; fixing one path from each latent variable to an manifest variable, there are 10 unknown such weights, 15 error variances, 2 residual error of prediction variances (for I and B_2), 6 variances and covariances of the shaded latent variables and 6 direct latent variable effects, making 39 parameters in all. The X^2 with 81 degrees of freedom (d.f.) ranged from 89 to 95 ($p>.05$ in all cases). In contrast, the Fishbein-Ajzen version of the model without B_1-B_2, B_1-I, and A-B_2 effects did not fit in any of the behavior domains; X^2s were 143-163 with 84 d.f. ($p<.05$). The 3 d.f. X^2 difference tests, of course, showed that the removal of the three paths yielded a significant decrement in fit of the model; thus, they are essential to an adequate understanding of the data in the context of the theories that were compared. Bentler & Speckart also provide a number of other model comparisons that need not be described here.

Test of a Theory of Intellectual Growth

Olsson and Bergman (1977) studied eight aptitude and achievement variables measured on two occasions at grades three and six, in a sample of 375 girls. They evaluated a model that proposed the eight variables could be represented by four primary intellectual latent variables on each of the two occasions of measurement, plus errors of measurement. They proposed that each latent variable would be regressed on itself across time, thus representing stability of the intellectual factors. They also allowed common factor residuals to be correlated in grade six, a feature not shown in their path diagram. This model fit the data quite well. Weeks (1978, 1979) proposed an alternative model based on the idea of "general intelligence." He hypothesized that each of the four primary intellectual factors could be further decomposed into a higher-order general factor plus a residual. Thus, general intelligence could affect performance on all eight measured

variables at each occasion, but only through the primary factors. Weeks' model is shown in Figure 4. In this model, manifest variables are given in squares, and all the remaining variables are latent variables. The independent variables in the system have been labeled $\xi_1 - \xi_{23}$, and all the dependent variables, including manifest variables and latent variables, as $\eta_1 - \eta_{27}$. Thus, there are 27 structural equations. Note that none of the independent variables are correlated; there are no two-headed arrows. The paths with numbers are fixed as known, in order to identify the model; the numbers were chosen for consistency with Olsson & Bergman's results. There are 53 free parameters to be estimated from the data.

The manifest variables $\eta_1 - \eta_8$ were measured in grade 3, while $\eta_9 - \eta_{16}$ represent the same manifest variables measured in grade 6. They represent Synonyms (η_1, η_9), Opposites (η_2, η_{10}), Achievement in Swedish (η_3, η_{11}), Letter groups (η_4, η_{12}), Figure sequences (η_5, η_{13}), Achievement in mathematics (η_6, η_{14}), Cube counting (η_7, η_{15}), and Metal folding (η_8, η_{16}) (See Olsson and Bergman 1977). The intercorrelations among these manifest variables are hypothesized to be generated by four first-order primary factors. Verbal comprehension (η_{17}), Inductive ability (η_{18}), Spatial ability (η_{19}), and a Knowledge or school achievement factor (η_{20}); each of these latent variables is presumed to be generated by general intelligence (ξ_2) and the first-order residual latent variables ξ_1 and ξ_3 (there are no residuals for factors η_{18} and η_{19}). Error of measurement latent variables $(\xi_8 - \xi_{15})$ are presumed to affect the manifest variables. A similar set of four primary factors $(\eta_{24} - \eta_{27})$, general intelligence (η_{22}), residual first-order factors (η_{21}, η_{23}), and errors of measurement $(\xi_{16} - \xi_{23})$ are presumed to exist in grade 6. In addition, η_{26} is not perfectly accounted for by general intelligence (η_{22}), leaving a residual (ξ_6); residuals ξ_4 and ξ_7 result from the regression of grade 6 on grade 3 residual latent variables; and general intelligence in grade 6 (η_{22}) is not perfectly predictable from its earlier counterpart ξ_2, yielding the residual ξ_5. Finally, errors of measurement in grade 3 $(\xi_8 - \xi_{15})$ are presumed to affect the corresponding manifest variables $(\eta_9 - \eta_{16})$ in grade 6, except for the $\xi_9 - \eta_{10}$ effect, which is taken as zero.

The importance of Weeks' model lies in the fact that it represents a pioneering attempt to define higher-order constructs in the context of causal modeling. While a mathematical basis for causal modeling with higher-order constructs was first given some years ago (P.M. Bentler, USPHS Grant DA 01070), Weeks (1978) provided the first statistical development of the approach which has now been further simplified (Bentler & Weeks, 1979a, 1979b). In this example, the general intelligence factor ξ_2 and its counterpart across time η_{22} represent latent variables that have no direct influence on manifest variables; their influence is indirect. As compared to the Olsson-Bergman model, in which the primary factors $\eta_{17} - \eta_{20}$ are simply allowed to correlate, in Weeks' model these correlations are themselves decomposed into a causal representation of a general influence ξ_2 plus a residual; a similar decomposition holds at the later time. The optimal parameter estimates obtained for Weeks' model were associated with a goodness of fit test that verified the plausibility of his theory. That is, the model could not be statistically rejected ($\underline{p} > .05$). Weeks (1979) provides details on the substantive meaning of various parameter estimates for the concepts of factor differentiation and stability.

Conclusion

Although there appear to be no published applications to experimental research, causal modeling has great potential in this area. By creating an appropriate model of possible effects in an experiment, and using indicators of hypothesized latent variables that are presumed to mediate the effects of an experiment, it can be determined whether the observed effects occur as hypothesized or in other ways. Causal modeling in experimental context has been discussed by various writers; see Bagozzi (1977), and Kenny (1979), for example; its role in clarifying univariate and multivariate analysis of variance has been spelled out by Rock.

Werts & Linn (1976), and by Rock, Werts & Flangie (1978); and its role in the analysis of covariance has been technically developed by Sörbom (1978).

The greatest contribution of causal modeling to psychology is liable to be in areas of quasi-experimental or nonexperimental research, where methods for testing theories are not well developed. Unfortunately, models based on manifest variables are more appropriate to problems of description and prediction than explanation and causal understanding: their parameters are unlikely to be invariant over various populations of interest. Thus, of particular importance, I believe, is the modeling of a process at the level of latent rather than measured variables, since the manifest variables only rarely correspond in a one-to-one fashion with the constructs of interest to the researcher, which will almost certainly be measured with error. As a consequence, conclusions about an manifest variable model cannot be relied upon, since various theoretical effects will of necessity be estimated in a biased manner. They also will not replicate in other studies that are identical except for the level of precision or error in the variables. Thus, the main virtues of latent variable models are their ability to separate error from meaningful effects and the associated parametric invariance obtainable under various circumstances. An example that compares latent variable and manifest variable models is given by Bentler and Huba (1979).

A few words are in order regarding terminology and new directions in the field. Latent variable models can be distinguished in the major traditional constrast of factor analytic vs. simultaneous equation models. The latter can be further divided as recursive vs. nonrecursive. Another distinction between models involves the contrast between models that ignore or account for the means of the manifest variables in the causal structure. A third distinction is concerned with the degree of abstractness of the latent variables and the generality and simplicity of the mathematical structure within which they are embedded.

In factor analysis, latent variables are related to each other

symmetrically. See Figure 2a and Bentler (1976a). There is no regression structure among latent variables; instead they are simply correlated or independent (oblique vs. orthogonal factors). In latent variable models with simultaneous equations, latent variables are also permitted to be regressed or predicted from other latent variables in various ways (see Figure 3). Consequently, multiple regression with independent variables subject to error is contained in these models, and more complex models have several such relations. There are two well known types of regression structures: recursive and nonrecursive. Somewhat paradoxically, nonrecursive structures are those that allow true "simultaneous" or reciprocal causation between latent variables, while recursive structures do not. Recursive structures have been favored as easier to interpret causally (Strotz & Wold, 1960). They generally are less difficult to estimate.

Another distinction between models involves the separability of first and second moment parameters. In separable models, the first moments (means) of the manifest variables are not structured in terms of the causal parameters. Consequently, they are effectively irrelevant to the causal modeling process and the goal is the modeling of the second moments (covariances, correlations, or cross-products) of the manifest variables. These models are thus often called covariance structure models (Bock & Bargmann, 1966), and they apply in most instances. More general moment structure models also allow an interdependence of first and second moment parameters (Bentler 1973, 1976b, Jöreskog, 1970; Sörbom 1974, 1978). The manifest variables' means are decomposed into basic parameters that may also affect the covariance structure. These models are particularly appropriate to studies of multiple populations or groups of subjects, and to the analysis of experimental data. They have not frequently been applied.

Linear structural equation models with latent variables have until recently been conceptualized as embodying only a first-order factor analytic measurement structure for the manifest variables. For example,

the JKW model expresses manifest variables directly in terms of latent variables. Thus, the latent variables are removed by one level from the manifest variables, as in Figure 3. However, it is also easy to conceptualize measurement models that are more complex, in which there are several levels or orders of latent variables, as in Figure 4. A general multilevel latent variable model that allows structured means was developed to deal with such situations (Bentler, 1976b). Such a model blurs the distinction made above between factor and simultaneous equation latent variable models, because the various levels of latent variables affect each other via regression structures. Weeks (1978) provided the first statistical development of a general model that allows multivariate regression structures on latent variables of various types and levels. Recent developments have been directed toward obtaining latent variable models that are both general and simple. These models allow causal influences across levels and types of latent variables (primary or residual latent variables at a given level), in addition to structured means. They generalize the JKW model (Bentler, 1979; Bentler & Weeks, 1979b; Lee & Bentler; 1979; Weeks, 1978). Bentler and Weeks (1979a) discuss these results and clarify the interrelations among a variety of structured linear causal models. An overview of statistical methods in structual equation modeling is provided by Bentler and Weeks (1979c).

Although the introductory nature of this chapter has limited the space that could be devoted to structural modeling in longitudinal contexts, it must be reiterated that causal modeling is particularly relevant to longitudinal research, since the passage of time helps to eliminate possible competing causal explanations of phenomena. Thus, it is not surprising that numerous methodological and empirical investigations exist in this area. See Bentler (1980) for relevant references, and Jöreskog (1979) and Rogosa (1979) for a further introduction to the field.

Literature Cited

Aigner, D.J. & Goldberger, A.S. (Eds.) Latent variables in socio-economic models. Amsterdam: North-Holland, 1977.

Ajzen, I., & Fishbein, M. Attitude-behavior relations: A theoretical analysis and review of empirical research. Psychological Bulletin, 1977, 84, 888-918.

Alker, H.R. Statistics and politics: The need for causal data analysis. In S.M. Lipset (Ed.), Politics and the social sciences. New York: Oxford, 1969.

Anderson, J.G. & Evans, F.B. Causal models in educational research: Recursive models. American Educational Research Journal, 1974, 11, 29-39.

Bagozzi, R.P. Structural equation models in experimental research. Journal of Marketing Research, 1977, 14, 209-26.

Bentler, P.M. Assessment of developmental factor change at the individual and group level. In J.R. Nesselroade & H.W. Reese, (Eds.). Life-span developmental psychology: Methodological issues. New York: Academic, 1973.

Bentler, P.M. Factor analysis. In P.M. Bentler, D.J. Lettieri & G. Austin (Eds.) Data analysis strategies and designs for substance abuse research. Washington: Government Printing Office, 1976. (a).

Bentler, P.M. Multistructure statistical model applies to factor analysis. Multivariate Behavioral Research, 1976. 11, 3-25, (b).

Bentler, P.M. The inderdependence of theory, methodology and empirical data: Causal modeling as an approach to construct validation. In D.B. Kandel (Ed.), Longitudinal research on drug use: Empirical findings and methodological issues. New York: Wiley, 1978.

Bentler, P.M., Linear simultaneous equation systems with multiple levels and types of latent variables. In K. G. Jöreskog & H. Wold, (Eds.), Systems under indirect observation, 1979. In press.

Bentler, P.M., Multivariate analysis with latent variables: causal modeling. Annual Review of Psychology, 1980, 31. In press.

Bentler, P.M., & Huba, G.J. Simple minitheories of love. Journal of Personality Social Psychology, 1979, 37, 124-130.

Bentler, P.M. & Peeler, W.H. Models of female orgasm. _Archives of Sexual Behavior_, 1979, _8_, 405-423.

Bentler, P.M., & Speckart, G. An evaluation of models of attitude-behavior relations. _Psychological Review_, 1979.

Bentler, P.M. & Weeks, D.G. Interrelations among models for the analysis of moment structures. _Multivariate Behavioral Research_, 1979, _14_, 169-185. (a)

Bentler, P.M. & Weeks, D.G. _Linear simultaneous equations with latent variables_. Los Angeles: University of California, 1979. (b)

Bentler, P.M. & Weeks, D.G. Multivariate analysis with latent variables. in P.R. Krishnaiah & L. Kanal (Eds.). _Handbook of Statistics, Vol. 2_, 1979, In press (c).

Bentler, P.M. & Woodward, J.A. Nonexperimental evaluation research: Contributions of causal modeling. In L. Datta & R. Perloff (Eds.), _Improving Evaluations_. Beverly Hills: Sage, 1979.

Bielby, W.T., & Hauser, R.M. Structural equation models. _Annual Review of Sociology_, 1977, _3_, 137-161.

Blalock, H.M. Correlation and causality: The multivariate case. _Social Forces_, 1961, _39_, 246-251.

Blalock, H.M. (Ed.) _Causal models in the social sciences_. Chicago: Aldine-Atherton, 1971.

Bloxom, B. Alternative approaches to factorial invariance. _Psychometrika_, 1972, _37_, 425-440.

Bock, R.D., & Bargmann, R.E. Analysis of covariance structures. _Psychometrika_, 1966, _31_, 507-534.

Boudon, R. A method of linear causal analysis: dependence analysis. _American Sociological Review_, 1965, _30_, 365-374

Campbell, D.T., & Stanley, J.C. Experimental and quasi-experimental designs for research on teaching. In N.L. Gage (Ed.), _Handbook of research on teaching_. New York: Rand-McNally, 1963.

Cooley, W.W. Explanatory observational studies. _Educational Researcher_, 1978, _7_, 9-15.

Cronbach, L.J. & Meehl, P.E. Construct validity in psychological tests. _Psychological Bulletin_, 1955, _52_, 281-302.

Duncan, O.D. Path analysis: Sociological examples. _American Journal of Sociology_, 1966, _72_, 1-16.

Fishbein, M., & Ajzen, I. _Belief, attitude, intention and behavior: An introduction to theory and research_. Reading, Mass.: Addison-Wesley, 1975.

Fishe, S. _The female orgasm_, New York: Basic Books, 1973.

Goldberger, A.S. Econometrics and psychometrics: A survey of communalities. _Psychometrika_, 1971, _36_, 83-107.

Goldberger, A.S. Structural equation methods in the social sciences. _Econometrica_, 1972, _40_, 979-1001.

Goldberger, A.S. & Duncan, O.D. (Eds.) _Structural equation models in the social sciences_, New York: Academic, 1973.

Jöreskog, K.G. Some contributions to maximum likelihood factor analysis. _Psychometrika_, 1967, _32_, 443-482.

Jöreskog, K.G. A general approach to confirmatory maximum likelihood factor analysis. _Psychometrika_, 1969, _34_, 183-202.

Jöreskog, K.G. A general method for analysis of covariance structures. _Biometrika_, 1970, _57_, 239-251.

Jöreskog, K.G. A general method for estimating a linear structural equation system. In A.S. Goldberger & O.D. Duncan (Eds.), _Structural equation models in the social sciences_. New York Academic, 1973 (a)

Jöreskog, K.G. Analysis of covariance structures. In P.R. Krishnaiah (Ed.), _Multivariate analysis-III_. New York: Academic, 1973 (b)

Jöreskog, K.G. Structural equation models in the social sciences: specification, estimation and testing. In P.R. Krishnaiah, (Ed.) _Applications of statistics_. Amsterdam: North-Holland, 1977.

Jöreskog, K.G. Structural analysis of covariance and correlation matrics. _Psychometrika_, 1978, _43_, 443-477.

Jöreskog, K.G. Statistical estimation of structural models in longitudinal-developmental investigation. In J.R. Neeselroade & P.B. Baltes (Eds.), _Longitudinal methodology in the study of behavior and human development_. New York: Academic, 1979, In press.

Jöreskog, K.G., & Sörbom, D. _LISREL IV Users Guide_. Chicago: National Educational Resources, 1978.

158

Keesling, W. Maximum likelihood approaches to causal flow analysis. Ph.D. thesis. Chicago: University of Chicago, 1972.

Kenny, D.A. Correlation and causality, New York: Wiley, 1979.

Leamer, E.E. Specification searches: Ad Hoc inference with nonexperimental data. New York: Wiley, 1978.

Lee, S.Y. & Bentler, P.M. Analysis of a general covariance structure model. Hong Kong: Chinese University, 1979.

Masters, W.H. & Johnson, V.E. Human sexual response. Boston: Little Brown, 1966.

Olsson, U., & Bergman, L.R. A longitudinal factor model for studying chance in ability structure. Multivariate Behavior Research, 1977, 12, 221-242.

Rock, D.A., Werts, C.E., & Linn, R.A. Structural equations as an aid in the interpretation of the non-orthogonal analysis of variance, Multivariate Behavioral Research, 1976, 11, 443-448.

Rock, D.A., Werts, C.E., Flaugher, R.L. The use of analysis of covariance structures for comparing the psychometric properties of multiple variables across populations. Multivariate Behavioral Research, 1978, 13, 403-418.

Rogosa, D. Causal models in longitudinal research. In J.L. Nesselroade & P.B. Baltes (Eds.), Longitudinal methodology in the study of behavior and human development. New York: Academic, 1979. In press.

Simon, H.A. Spurious correlation: A causal interpretation. Journal of The American Statistical Association, 1954, 49, 467-479.

Singer, J. & Singer, I. Types of female orgasm. Journal of Sex Research, 1972, 8, 255-267.

Sörbom, D. A general method for studying differences in factor means and factor structure between groups. British Journal of Mathematical and Statistical Psychology, 1974, 27, 229-239.

Sörbom, D. Detection of correlated errors in longitudinal data. British Journal of Mathematical Statistical Psychology, 1975, 28, 138-151.

Sörbom, D. An alternative to the methodology for analysis of covariance. Psychometrika, 1978, 43, 381-396.

Sörbom, D. & Jöreskog, K.G. COFAMM, confirmatory factor analysis with model modification. Chicago: National Educational Resources, 1976.

159

Strotz, R.H. & Wold, H.O.A. Recursive vs. nonrecursive systems: An attempt at synthesis. Econometrica, 1960, 28, 417-427.

Torgerson, W.S. Theory and methods of scaling. New York: Wiley, 1958.

Tucker, L.R., & Lewis, C. A reliability coefficient for maximum likelihood factor analysis. Psychometrika, 1973, 38, 1-10.

Weeks, D.G. Structural equation systems on latent variables within a second-order measurement mode. Ph.D. thesis. Los Angeles: University of California, 1978.

Weeks, D.G. A second-order longitudinal model of ability structure. Multivariate Behavioral Research, 1979.

Werts, C.E. & Linn, R.L. Path analysis: Psychological examples. Psychological Bulletin, 1970, 74, 193-212.

Wiley, D.E. The identification problem for structural equation models with unmeasured variables. In A.S. Goldberger & O.D. Duncan (Eds.), Structural equation models in the social sciences. New York: Academic, 1973.

Wright, S. The method of path coefficients. Annals of Mathematical Statistics, 1934, 5, 161-215.

Figure Legends

<u>Figure 1</u>. Two representations for a structural equation: (a) path diagram for the equation; (b) the complete model.

<u>Figure 2</u>. Latent and manifest variable models: (a) a factor analytic model; (b) a path model.

<u>Figure 3</u>. A model for attitude-behavior relations.

<u>Figure 4</u>. A model for adolescent intellectual growth.

Figure 1

(a)

(b)

Figure 2

(a)

(b)

162
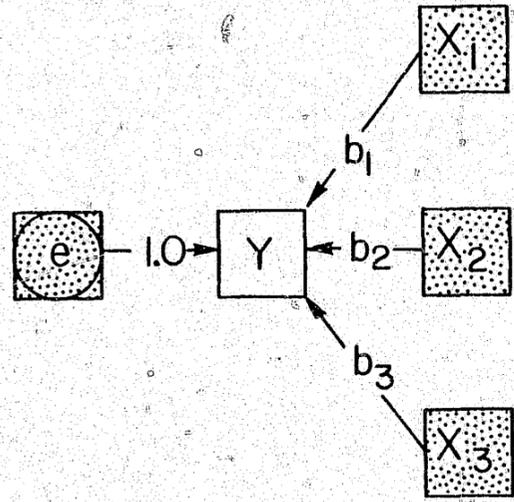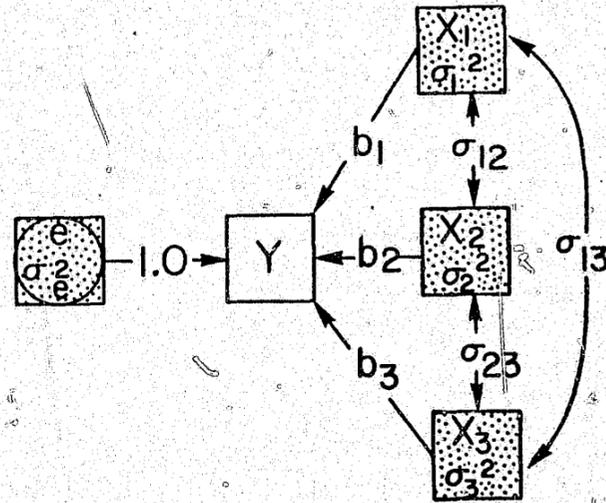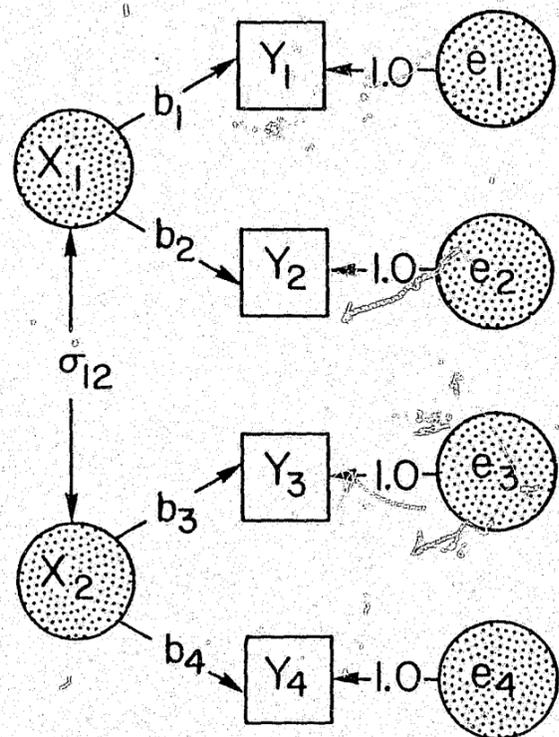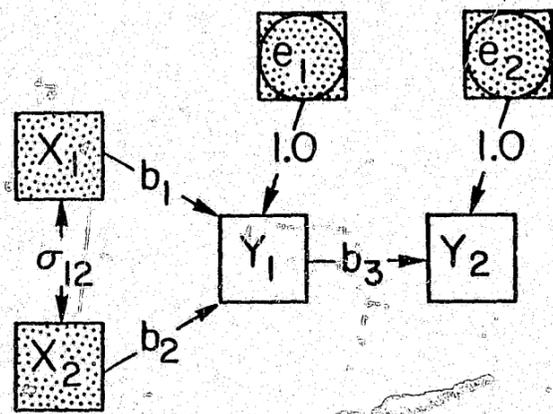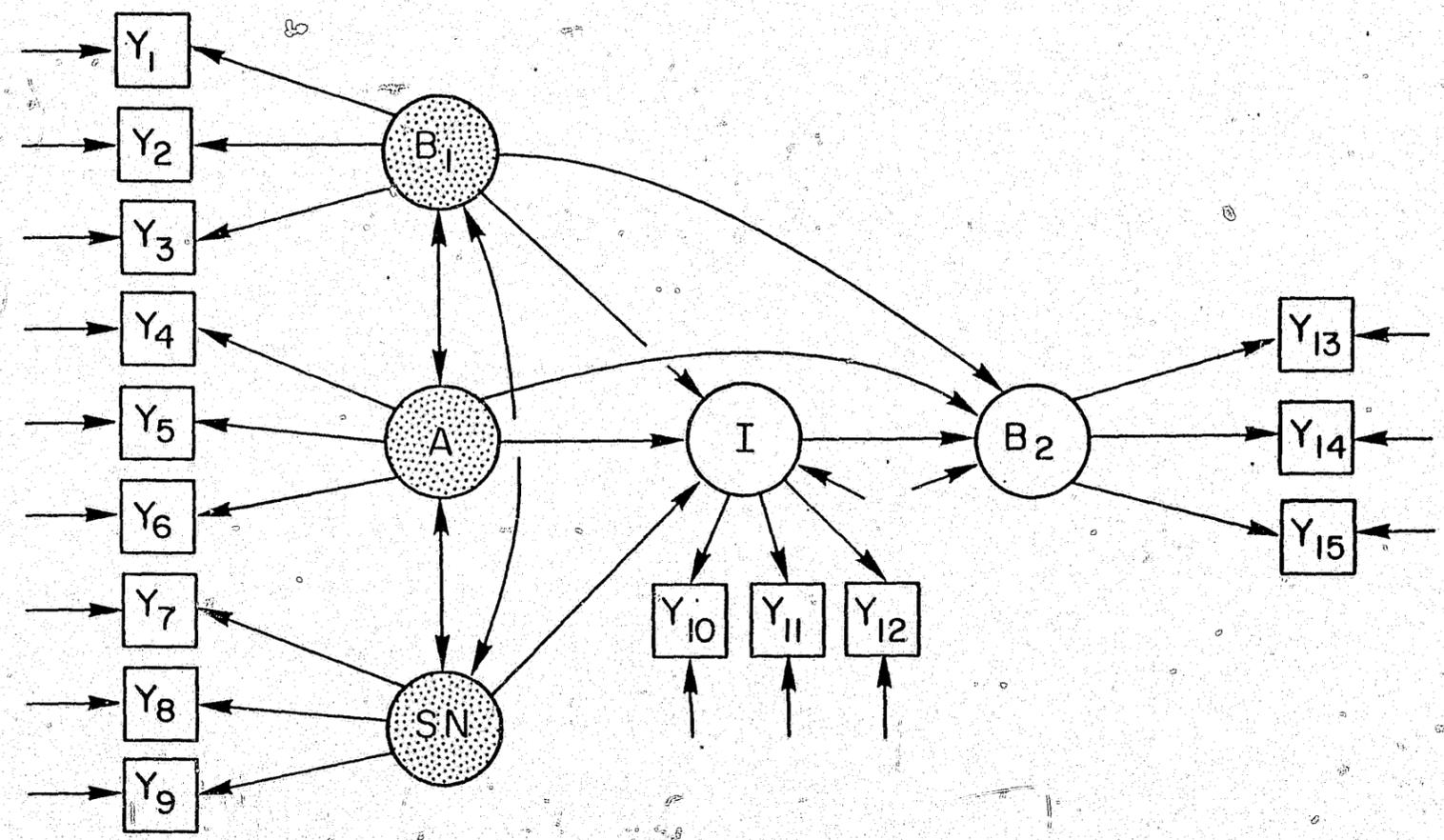
Figure 3

Figure 4
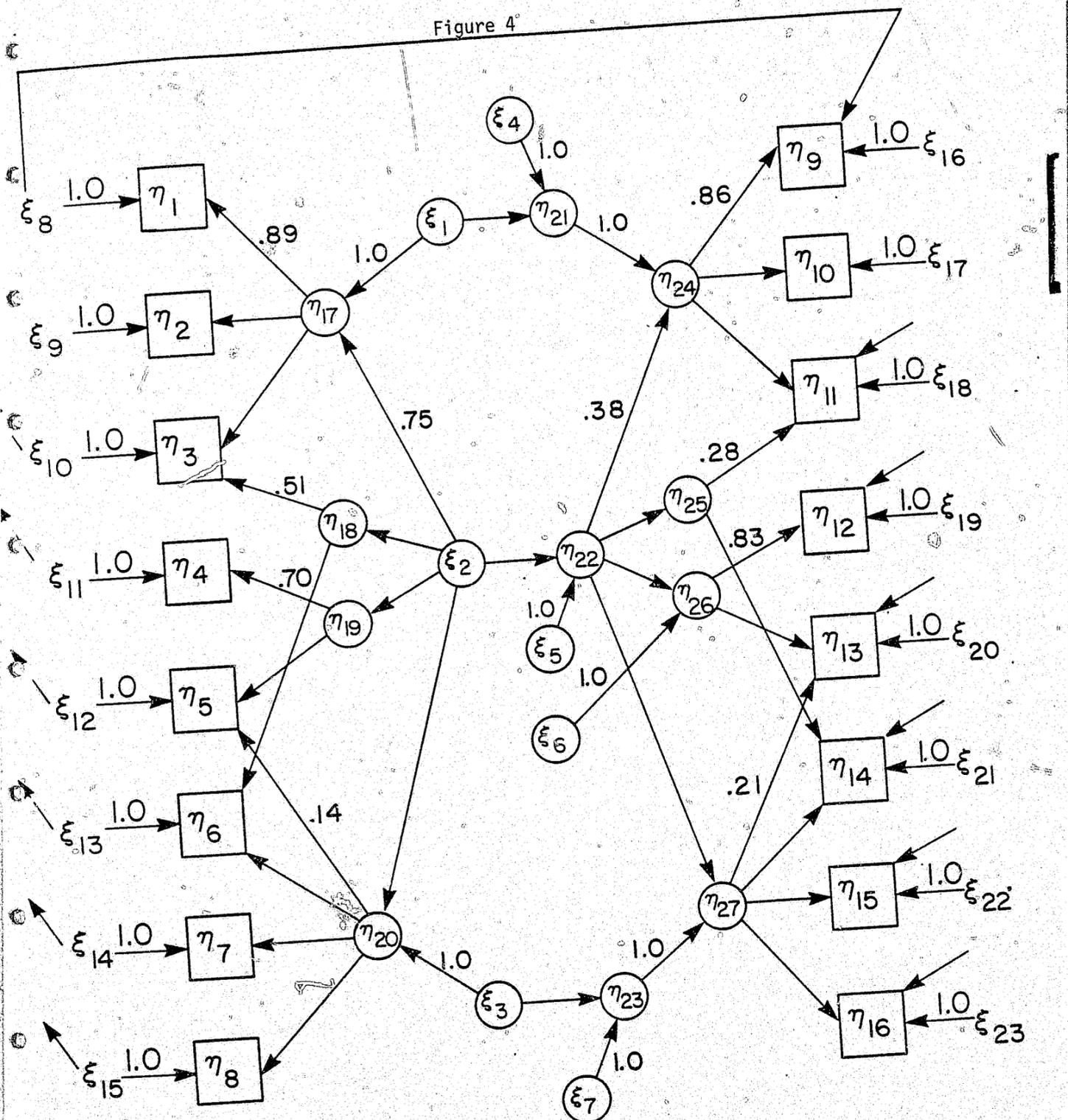
Chapter 7

An Application of Causal Modeling Techniques

to Prospective Longitudinal Data Bases[*]

Robert L. Baker

Birgitte R. Mednick

University of Southern California

William Brock

SWRL Educational Research and Development

An Application of Causal Modeling Techniques
to Prospective Longitudinal Data Bases

Robert L. Baker

Birgitte R. Mednick

University of Southern California

William Brock

SWRL Educational Research and Development

Introduction

This paper will describe how exploratory path analysis may be used to draw causal inferences between and among social antecedent and medical outcome variables measured over a 21 month period (from the beginning of pregnancy to age one year). The data to be used in this analysis stem from a Danish birth cohort consisting of 9125 deliveries. The analyses constitute a part of a comprehensive study following up this Danish sample to age one year. The purpose of this study is to examine the impact of a set of social, familial and maternal state variables on the medical status of neonates and one year olds. In addition to the above-mentioned predictors, a series of variables describing the medical conditions of the subjects at different points in time are included. The function of these variables in the analyses is primarily to provide medical outcome data possessing different temporal and clinical characteristics against which to evaluate the relative predictive strength of the non-medical variables. Thus, in the data interpretation, the interrelationships observed between the purely medical variables will receive only limited attention.

Prediction of Outcomes in Early Childhood

Perinatal and Neonatal Antecedents

A large amount of published research has dealt with prediction of outcomes within the first year of life on the basis of measures of perinatal and neonatal complications or deviance (Parmelee, et al., 1976; Prechtl, 1965; Rosenblith, 1966; St. Claire, 1978; Parmelee and Michaelis, 1971; Werner, et al., 1971; Goldstein, et al., 1976).

The most frequently used outcome variables in this body of literature are measures of cognitive and neurological/motor developmental functioning (Niswander, et al., 1966; Drage and Berendes, 1966; Dijkstra, 1960; Honzik, et al., 1965; Field, et al., 1978; Rubin, et al., 1973; Goldstein, et al., 1976). Only a few studies have studied the physical health of the young child as an outcome variable in this area of research (Werner, et al., 1971; Fitzhardinge and Ramsey, 1973).

Sameroff (1979) summarized this body of literature. The overall finding is that perinatal and neonatal problems appear to be associated with increased neurological, physical and cognitive problems during the pre-school years, but that this effect ceases to be significant at the time of school entry.

Environmental and Maternal State Antecedents

Variables describing the developmental environment clearly constitute another important source of influence on the development during early childhood. Our knowledge of the effects of environmental variables, e.g., SES, stems primarily from studies of cognitive development (Sameroff, 1979; Bradley, et al., 1979; Bronfenbrenner, 1974). The overall pattern of results suggests that environmental factors have no measurable effect on cognitive development during the first year of life. Sometime during the second year, most often around age 18 months, the negative effects of less optimal environmental conditions begin to appear. The exceptions to this general finding are noted in those instances where specific environmental changes were established as part of intensive preventive intervention program (Scarr-Salapatek and Williams, 1973; Heber, 1975). In such cases, positive effects of the systematic intensive manipulation of the environment were observed within the first year.

Implications for the Present Study

The literature on prediction of outcomes during early childhood consistently shows that outcomes during the infant's first year are

most difficult to predict. The effects of perinatal complications are
far from dramatic and the impact of SES and other environmental status
variables may not yet be observed. The question of why there is such
a low level of sensitivity has been analyzed by several authors
(Sameroff, 1978; Crano, 1977). Among the most frequently cited explana-
tions, particularly related to cognitive and neurological outcomes, are
the qualitative discontinuity of what are defined as measures of cognitive
competence during early childhood (Sameroff, 1979) and the fact that the
measures of cognitive as well as neurological functioning obtainable
during the period are too gross to reflect the more subtle individual
deviations (Rubin, et al., 1973; Mednick, 1978). The contribution
of this causal analysis of one year infant outcomes to the extant
literature rests with the fact that the analyses reported here assess
the effects of environmental variables on three types of medical one
year outcomes. Whereas cognitive measures have not reflected an impact
by environmental variables at this age, it is possible that physical
health and motor development measures, being less qualitatively dis-
continuous, may be more sensitive to variations in environmental
variables at an earlier point in time or age. The present study offers
an opportunity to examine the validity of this notion. If it can be
shown that environmental variables significantly impact any of the
three medical measures at age one, in addition to its theoretical im-
portance, it will have heuristic value for the design of preventive
intervention procedures for use with special populations of children.

Description of the Danish Cohort

The subjects of the study consist of the total birth cohort
included in a Danish prospective longitudinal perinatal project. The
project was begun in 1959 at the maternity department of the State
University Hospital (Rigshospitalet) in Copenhagen.[1] All deliveries
(over 20 weeks gestation) that took place in this hospital between
September 1959 and December 1961 were included in the study. The
following data collection procedures were employed by the original

---

[1]The Danish Perinatal Study was carried out under the guidance of
Professor P. Plum and D. Trolle. In addition, Professors E. Rydberg,
E. Brandstrup and F. Fuchs gave expertise and support in the planning
phase as well as during the data collection phase. Professor B.
Zachau-Christiansen and A. Villumsen, M. D. served as project directors
and also carried out the medical examinations of the pregnant women and
their infants.

project collaborators (Zachau-Christiansen and Ross, 1975): The
pregnant women were contacted and examined before delivery; if pos-
sible, very early in pregnancy, during attendance at the hospital's
antenatal clinic. To evaluate and code the social, general medical
and obstetric histories of the women uniformly, all prenatal examina-
tions were done by the same physician. In addition to an obstetrician,
midwives and midwife trainees were present in the delivery room and
assisted in collecting the data describing the deliveries and the status
of the neonates and mothers immediately after birth. In those cases
where the general condition permitted, the live-born infants were again
examined on the first and fifth day after delivery by one of three
pediatricians. The first and the fifth day examinations included a
physical examination and a thorough neurological assessment of the
infant. Upon discharge from the hospital the mothers (or guardians)
of the infants received a self-administered questionnaire related to
the infants' developmental progress during the first year of life. In-
formation concerning attendance at the free national infant health
examination, intercurrent diseases, admission to hospitals and other
institutions, and records of immunizations were also obtained. When
the children reached their first birthday, the mothers were asked to
bring them to the pediatric out-patient department of the State Univer-
sity Hospital for a special developmental examination. A team of
three pediatricians conducted the follow-up examinations for all of the
surviving infants. Maximum effort was made to include all surviving
infants in the one year follow-up. In cases where the parents were not
able to bring the child to the hospital, home visits were arranged.
The one year examination included a full neurological assessment, an
evaluation of the motor developmental level and of the physical health
status at the time of the examination.

All the information collected, i.e., the anamnesis of the preg-
nant women and the descriptions of the pregnancies, deliveries, and
neonatal and one year status of the infants, were pre-coded.

Reduction of Medical Data Items into Composite Scores

For the purposes of this study, the many individual data items
from the subjects' medical protocols were summarized into sets of
composite scores. These scores are labeled:

   o pregnancy complications

   o delivery complications

   o neonatal physical status

   o neonatal neurological status

o one year physical status

o one year neurological status

o one year motor developmental
  status.

The principles governing the organization of the many hundred individual data items into these summary scores were developed through the collaboration of American and Danish obstetricians and pediatric neurologists. After serious attempts at factor and cluster analyses had not proven feasible, a system was developed that assigned a weight (0-5) to each data item according to the judged severity of complication or deviance. Following are the kinds of scores developed on the basis of this system: 1) the summed weight score--indicates the sum of all the weights assigned to the complications within a given area; 2) the highest severity--indicates the highest weight assigned to any of the problems observed within the area; and 3) the frequency score--simply indicates the total number of complications within an area. These composite medical scores have been used in previous publications (S. Mednick, et al., 1971; B. Mednick, 1977). In the present analyses the summed weighted scores have been used.

Characteristics of the Danish Sample

If explanation is the research outcome sought rather than prediction alone, than those special characteristics of the data base and research questions that might be influential in constructing the causal model and drawing causal inferences should be identified in advance.

Representativeness

The Danish perinatal sample is not a representative sample of pregnant Danish women. The differences between this sample and the general Danish population of prospective mothers are related to: 1) medical risk; 2) social background of subjects; and 3) the medical treatment to which the subjects were exposed.

Medical Risks

The Danish perinatal sample contains a higher than normal rate of problem pregnancies and deliveries. The State University Hospital, Rigshospitalet, is a unit of the University of Copenhagen

and is also the country's largest center for medical research and progressive treatment of patients. It is, therefore, not surprising that a larger percentage of difficult births are referred to this hospital. During the period 1959-1961, the hospital system of Denmark was still so centralized that women from all the eastern parts of the country who were suffering grave complications in pregnancy, or who were expected to present difficult deliveries would likely be referred to Rigshospitalet's obstetric department. As an example of the elevated complication rate, the incidence of prematurity (i.e., birthweight under 2500 grams) was over 18% in the Danish perinatal sample, or about three times the incidence found in the general Danish population during the years in question.

Social Background

Women in the Danish perinatal sample experienced a higher rate of unwed motherhood and a lower average SES compared with representative samples of Danish pregnant women. This is likely due to the fact that Rigshospitalet is located in the center of the city and therefore draws a sizeable proportion of the patient group not referred because of complications from the poorer inner-city areas. There was also a traditional tendency during those years to have unwed mothers from Copenhagen give birth at Rigshospitalet.

Medical Treatment

The quality of treatment received by pregnant women in the general population is known to vary as a function of such variables as SES, age of mother, area of residence, degree of wantedness of pregnancy, etc. (Kessner, et al., 1973; Chamberlain, et al., 1975; Pharoah, 1976). However, these factors did not influence the quality of treatment to which the subjects of the Danish perinatal sample had access. Irrespective of social or personal background, prospective mothers received early and thorough prenatal care. During delivery, as well as during the postnatal period, all patients in the study sample were attended by highly trained medical personnel using the most advanced equipment available for the treatment of both mother and infant.

Comparability with Other Samples

While the described sample characteristics render the Danish perinatal sample inappropriate for the calculation of incidence rates generalizable to the total Danish population, the data from the sample present unusual opportunities for analysis of relationship between pre- and perinatal conditions and later outcomes in well-defined subject

groups, i.e., groups defined in terms of social as well as medical variables. The white subsample of the American Collaborative Project constitutes an analogous American sample, i.e., this sample presents higher incidences of young, low SES women and women presenting higher degrees of medical risk in comparison to representative U. S. samples (Niswander and Gordon, 1972). The comparability of these two studies and specifically the relevance of findings from the Danish perinatal sample to similar American samples has been demonstrated (Mednick, Baker and Sutton-Smith, 1979).

Description of the Danish Study

Since the analyses described here are embedded in a larger study, a short description of the comprehensive study plan is presented.

The study has two major parts. The first part emphasizes the peri- and neonatal outcomes of the total birth cohort. The second part focuses on the one year status of the infants. Table 1 presents the variables included in the analyses relevant to the neonatal level. The variables are grouped into three categories: antecedent, mediating, and outcome variables. The categories were designed on the basis of what appeared to be the most likely temporal sequence of the variables' impact on the mother-child dyad. The category of antecedent variables includes social, demographic, and family structure data which were determined at the time of conception. In addition, medical data describing the mother's pre-pregnancy health and previous reproductive history are included in this category. The effects (or association) of these antecedent variables on the outcome variables are seen as mediated by pregnancy and delivery complications. The outcome variables analyzed include birthweight, and two summary scores describing the physical and neurological status of the newborn.

Table 2 shows the variables included in the second part of the study, i.e., the analyses of the one year outcomes. The organization of the variables into the three overall categories was done according to the same principles as that described for Table 1. In addition to the variables described in Table 1, Table 2 includes a series of variables reflecting the social environment of the child during the first year, i.e., the two subcategories of mediating variables labeled sociological factors and parental health. It should be noted that the variables listed as outcome variables in Table 1, i.e., the neonatal outcomes, are considered antecedent variables in Table 2. In other words, in the assessment of developmental progress over the first year, the pre-, peri-, and neonatal status of the infant were taken into account.

Development of the Path Models

Path analysis is a multivariate correlational procedure originally identified with biometrics (Wright, 1934). In recent years, there has been an increased interest in such procedures applied to the social sciences (Bentler, 1980).

The procedural framework which has been utilized in the present inquiry is that of Joreskog and Sorbom (1978). They developed a model for the analysis of linear structural equations via a maximum likelihood technique. The computer program (LISREL IV) is capable of handling both measured and unmeasured (or latent) variables and is sufficiently versatile to permit some degree of model manipulation to improve the 'fit' of a model to the data. This capability allows exploratory model building and testing strategies to be superimposed upon what is essentially a confirmatory factor analytic procedure. Application of path analytic procedures involves initial development of a comprehensive generalized model to be tested by the available data. Ideally, this model should be developed on the basis of extant theories in the area of inquiry. While the published literature on the social, familial, and maternal state variables is far from scant, its use in suggesting a generalized causal model is limited. Thus, the architecture of such a model must be generated in large part by the inquiry itself, i.e., through a sequence of preceding analyses leading to the formulation of the final model. In this endeavor we are helped by the fact that in spite of its lack of comprehensive theories, the literature in the area is rich in suggestions of networks of testable hypotheses. Since the emphasis in this paper is on the application of the path analysis technique, we will briefly describe the general sequence of analyses that led to the choice of a causal model.

The five stage progression was not intended as a general iterative paradigm. However, each stage did constitute a sequential (but not linear) progression of activities useful in reducing the myriad of variable and temporal relationships to a set of useful functional relationships.

Crosstabular Analyses

The initial step of the data analysis for this study included a series of crosstabular analyses of the distributions of the main antecedent and mediating variables. Recognizing the limitations of such an analytic technique (Kerlinger and Pedhazur, 1973), it was used here as a set of preliminary "ordering" analyses.

Exploratory Factor Analyses

Factor analytic techniques were used for data reduction and preliminary hypothesis formation purposes. In addition to identification of the resulting factor structure(s), special attention was given to determining the factorial complexity of medical and other outcome variables included. The purpose of this step was to dimensionalize the data base, thereby initiating the reduction of the number of data items that must be dealt with.

Hierarchical Multiple Regression Analyses

An array of social and medical variables was entered into a series of step-by-step multiple regression analyses of the total sample predicting toward each of the five medical outcomes at the neonatal and one year levels. In an attempt to infer causation, selected medical and social variables were included for estimating interactions. Although not thoroughly accommodated by the path model, the regression analyses of the total population of 9125 placed us in position to better understand the independent contributions of each of the variables. At the same time, the interactions began to suggest possible causal relationships.

Discriminant Function Analyses: A Temporal Composite

In the regression analyses the variance associated with critical subpopulations is confounded with the variance associated with the total sample. It was desirable to isolate defined subpopulations and obtain maximum separation between them on the basis of scores on selected variables or on patterns of scores and configurations of several variables hypothesized to be "critical" predictors. As part of this analytic activity, seven discriminant function analyses were completed, one for each of the medical outcome variables. Two criterion groups were identified for each of the analyses on the basis of the weighted scores on the dependent variable; a group of cases with very bad scores and a group with very good (i.e., normal scores) constituted the two groups for each analysis. Our interest was in determining the contribution of selected medical and social antecedent variables to the prediction of positive and/or negative scores on medical outcome variables. In most cases, definitions of the subsamples in effect reduced the variance on selected comparison variables to near zero. Thus, relationships that would otherwise be masked due to confounding with total sample variance would become a great deal more apparent in such comparisons.

The analyses were graphically arrayed to provide suggestions of causal and temporal influences of the variables. The seven medical outcome variables occur in the four-stage temporal frame shown for the "Y"

variables in Figure 1. A total of 18 predictors (or what are called "X" variables in Figure 1) were available for inclusion in the analyses. However, not all 18 were included in every analysis due to such considerations as temporal irrelevance. In addition to the inclusion of the common set of predictors in each analysis, each subsequent analysis also included all of the medical outcome variables from the previous analyses as predictor variables. Figure 2 shows a temporal composite of the seven discriminant analyses. Remember, each of the seven analyses was comprised of a different sample of the cohort. That is, those mothers in the "high" group in the pregnancy complication variable analysis were not the same (necessarily) as those in the "high" group for delivery complications.

The figure and schema used to "lash" together the seven analyses to include the "outcome-over-time" interaction were used here simply to approximate some causal paths. Thus, they are intended only as instructive speculation. However, it seemed defensible on the basis of this speculation and of the existing literature to develop a causal model which may then be further analyzed via exploratory path analysis procedures.

Generalized Model for Outcomes During the First Year

An initial generalized causal model was developed on the basis of these previous analyses and extant literature. Figure 3 presents the model. Each square box indicates a measured (observed) variable. Because all variables in this model are measured this is an observed model. Table 3 presents the full variable names for all variables in the model. Double-headed arrows in the system indicate that two variables are correlated. A uni-directional arrow indicates that there exists a one-way 'causal' relationship between the two variables. All the variables in the left column are considered to be inter-correlated and will be referred to as independent or X variables. In the two final models, double-headed arrows will be drawn only where a correlation exists that is significant at the .05 level. Although this notation greatly simplifies the pictorial version of the model, it is important to keep in mind that all variables in the left column correlate freely. In the initial model the unidirectional arrows which lead to the dependent variables without a causative variable indicate the 'residual' variance of the dependent variables. Highly correlated independent variables (exogenous) make it difficult to identify causal inferences since path estimates can differ dramatically from one sample to another (Asher, 1976; Deegan, 1972). However, since multicollinearity is a problem arising from the instability of sample estimates, maybe the Danish cohort is not so much at risk. Deegan (1972) has shown that large sample size can reduce the impact of collinearity. Not only do we have adequate sample size, we are also able to draw multiple samples

for "cross validation" purposes.

In complex models such as the one presented here, the distinction between independent and dependent variables frequently is blurred. In the present models the Y variables (pregnancy complications, delivery complications, birthweight, neonatal neurological and physical scores, and the first year outcome measures) are firmly embedded in a known time frame (see Figure 3). The X variables have an undetermined time of impact and, in many cases, probably have ongoing impact. For convenience, all variables describing pre-pregnancy medical conditions are considered "X" variables even though their temporal characteristics may be known and invariant, e.g., negative data on previous pregnancies, or pre-pregnancy health. In a model based on a large-scale observational data base such as the present study, the residual variance will likely be quite large, while the "explained" variance will be relatively small. This is primarily due to three factors: 1) the relative homogeneity of the group with respect to initial prenatal care, 2) the extreme heterogeneity of the sample population on all other variables, including social influences and 3) effects of distribution.

Model Testing and Fitting Procedures

The generalized model was applied independently to each of two random samples (N=400) for which models were to be tested and modified. The two samples used in the path analyses are missing data correlation matrices as defined by Cohen and Cohen (1975) utilizing pair-wise deletion of missing data (minimum n=349).

Tests for goodness of fit. Several statistics and measures have been utilized to describe the extent to which the specified model is an accurate descriptor for the input data (see McGarvey, 1977). In this study, the measures of goodness of fit were an overall chi square test (testing the goodness of fit for the entire model) and the standard errors of individual parameter specifications (for the significance of each coefficient). The closer the chi square value is to the degrees of freedom for the model the better the indication of a good fit, i.e., the ratio of the X^2 value to the degrees of freedom should be minimized. The path coefficients should be significantly different from zero (typically measured at the .05 level in this study).

Model modification: the elimination of nonsignificant coefficients. An examination of the standard errors of the coefficient reveals those which are statistically non-significant and are considered to detract from the overall fit of the model. In subsequent tests of the model it is possible to drop the nonsignificant parameters in an effort to improve the overall model.

Model modification: the addition of new coefficients to be estimated. Two primary methods have been suggested for the rational addition of new paths to models, both involving information derived from the LISREL solution. The residuals of the equation can be examined to determine the pattern of data poorly fitted to the model (Bentler, 1980) and the first order derivatives of the equation may also provide information about specific patterns of lack of fit (Sorbom, 1975).

Model modifications were undertaken cautiously since there are a number of pitfalls possible in such a post-hoc procedure. When modifying a model to improve its fit to the study sample, paths may be added or deleted. The elimination of paths increases the degrees of freedom for the model. The addition of paths may decrease the overall X^2 when that addition improves the fit of the model. Thus, adding and deleting paths can contribute to minimization of the X^2/df ratio, which is the best overall measure of an improved model.

The development of two models permits some comparisons of variable relationships across models. The path coefficients computed in these analyses are unstandardized. Asher (1976) indicates that while it is common sense and not formal rules that provides the best guides to determining whether to use standardized or unstandardized estimates, he does advise that when one wishes to make comparisons across subsets of data unstandardized coefficients are appropriate since they are immune to the effects of the different variances in the same variable that may arise due to subsetting. Heise (1975) points out that due to its dependence on population distributions, standardization is to be avoided in comparative studies of different populations or in longitudinal studies of a changing population. The above is academic if one uses the correlation matrix as opposed to the variance-covariance matrix as the basis for the analysis, since the coefficients are actually standardized as a function of collapsing the variance.

Results of the Analyses

The original generalized model was tested and modified to obtain the best fit with the actual data from two independent samples that were randomly selected from the total Danish birth cohort. Figures 4 and 5 present the path diagrams and unstandardized coefficients for the models produced by the two samples. Figure 6 presents the model showing only the paths that are significant in both models. A "loose" procedure (after Bentler, 1980) was used in that identical model and fitting procedures were applied to both samples. Thus, the statistically significant path coefficients common to the final models of both samples will represent a high probability of being more than chance findings. Paths that are statistically significant in only one sample carry a greater risk of

being products of error variance specific to that sample. For this reason, the focus of the result will be on the path coefficients which are significant in both of the final models. In one case, however, an exception is made to this rule because the path in question is specifically relevant to an issue in the existing literature.

Intercorrelation Between the Exogenous (X) Variables

The correlations discussed here between the exogenous variables were found to be significant in both samples. Table 4 presents the correlation matrices for the two samples. Sample 1 correlations are above the diagonal and Sample 2 correlations are below. Use of daycare was naturally related to mother's employment. The magnitude of the relationship in the two samples appears smaller than expected (.36 and .49). However, daycare in these analyses only incorporates information about institutional daycare. Previous analyses have shown that exposure to private daycare arrangements yielded no relationship with any of the outcome variables studied.

Mother's age was correlated with family size, SES, and number of years between the index child and the previously-born child. That is, older mothers tend to have: larger families, a longer time since the previous childbirth, and a higher SES classification. The reason that mother's age and SES correlate negatively is that the British system used for SES classification (Classification of Occupations, 1965) is constructed so that a higher code corresponds to lower SES.

Significant relationships were found between SES and wantedness of the pregnancy (r's= .33 and .37). Higher SES mothers seem to want the pregnancy more (wantedness is scored so that a higher score is associated with not wanting the child).

Patterns of influence on the endogenous (Y) variables pregnancy complications. Only one of the variables included showed a direct relationship with the variable indicating the composite amount of pregnancy complications observed. The longer the spacing from the birth of the previous child, the higher the pregnancy complications score. In addition, an indirect relationship between the mother's age and pregnancy complications (via spacing) occurs in both samples. In previous analyses the variable mother's age has shown a significant independent effect on pregnancy complications. In the path analyses, spacing and mother's age are highly correlated and spacing shows a slightly stronger correlation with pregnancy complications than does mother's age and prevents mother's age from manifesting anything but an indirect relationship with this outcome variable.

Delivery complications. Complications of the pregnancy and previous health problems in the mother showed significant direct effects on amount of delivery complications observed in the subjects. In addition, the variable of delivery complications was influenced by the indirect effects of mother's age and spacing via pregnancy complications.

Birthweight. Only one of the variables included in this model was found to have a significant impact on birthweight. In both samples, multiple births was related to lower birthweight.

Neonatal physical status. Birthweight, delivery complications and family size showed significant effects on neonatal physical status. It appears that family size must be interpreted as a biological variable rather than an environmental variable. Actually, family size must be interpreted as an indicator of physiological ability to bear additional children as a result of having borne several other children. The social effects of family size have not had sufficient time to become manifest by the end of one year. Family size is strongly related to mother's age, and the indirect effect of mother's age is partly responsible for the family size effect. The reason family size, rather than mother's age, shows a direct effect on the neonatal physical status of the infant, may be found in an argument similar to that presented above for spacing and mother's age. Neonatal physical status is influenced indirectly by the network of variables that effects the three variables that do impact directly on this outcome. It should also be noted that the correlation between the two concurrently obtained outcome measures, i.e., neonatal physical and neonatal neurological status, is very high (r=.37 and .35).

Neonatal neurological status. This outcome variable is directly influenced by birthweight and delivery complications; and it is indirectly effected by those variables that directly and indirectly influence each of them.

One year physical status. This outcome, indicating the physical health status of the child during the first year, is directly influenced by birthweight, neonatal neurological status, and indirectly by neonatal physical status (these two concurrently obtained measures were allowed to be correlated freely in this model). In addition, exposure to daycare significantly impacts on the one year health status and is indirectly effected via daycare by mother's employment. As is the case with the previously discussed outcomes, the pattern of direct and indirect effects influencing the three variables that directly influence one year health status can also be interpreted as having indirect effects on one year physical status.

One year motor development. Birthweight is the only variable which significantly influences this variable in both samples.

<u>One year neurological status</u>. No variable significantly influences this outcome in both samples. Whereas the other two one-year outcomes, i.e., physical health and motor development, are highly correlated in both samples, one year neurological status is correlated with physical health status in both samples but with motor development in only one. Of the three one year outcomes, one year neurological status is clearly the one we are least able to predict using the variables and measures available in the described model.

Data Interpretation

Since the focus in this paper is on the effects of a series of social and non-medical variables included in the analyses on the one-year measures, only limited attention will be given to the inter-relationships between the purely medical variables.

<u>Neonatal status</u>. The data analyses pertaining to the neonatal outcomes, i.e., birthweight, neonatal physical and neonatal neurological status, indicate that the included non-medical variables have no effect on neonatal outcomes in the Danish perinatal sample. The direct and indirect relationships between family size, mother's age and neonatal physical status appear to be mediated by biological covariates of these independent variables rather than by any of the non-medical or envionmental covariates. This interpretation is supported by the well replicated finding that increases in mother's age and the highly correlated variable family size are related to an increased level of reproductive problems (Niswander and Gordon, 1972; Butler & Alberman, 1969). This relationship has been shown primarily to be caused by biological changes in childbearing as a function of these factors (Mednick, Baker & Sutton-Smith, 1979) rather than by older women or women with large families being disproportionately represented in environments conducive to reproductive problems. An interpretation centering on biological mediation of the effect of family size is further supported by the fact that other environmental variables such as SES and wantedness of pregnancy (with one exception in one sample) showed no impact on any of the pre- or neonatal outcomes. This finding is in conflict with the generally reported finding described in the introduction of a significant relationship between social variables (such as SES) and pregnancy outcome.

Two studies reported in the literature present findings comparable to those of the Danish perinatal sample, i.e., no environmental influence on neonatal outcomes. In a study of a complete Hawaiian birth cohort (Werner, et al., 1971), including all pregnancies occurring on the island of Kauai during a given time period, social variables such as SES and family stability were found to have no impact on peri- and neonatal outcomes. In the white subsample of the American Collaborative Perinatal Project (Niswander & Gordon, 1972) unwed mothers and teenage mothers

showed lower incidences of perinatal mortality and complications than did other categories of women, e.g., married women in their twenties. This study did not analyze perinatal outcomes as a function of SES; however, the variable mother's education was analyzed and did not show any systematic effect on perinatal outcomes. The analyses presented in the paper presented here do not provide any evidence regarding the effects of teenage motherhood and illegitimacy in the Danish sample. However, a previous publication (Mednick, Baker and Sutton-Smith, 1979) found that a pattern of results similar to that reported for the American Collaborative Study was observed in the Danish data.

The most probable explanation for the lack of effect of environmental variables on peri- and neonatal outcomes in these three studies may be that the samples in these studies in every case received through medical care irrespective of SES and other social factors, starting early in pregnancy and lasting throughout the post-delivery period. In the general population, on the other hand, social factors are known to influence the quality of the medical care available as well as the importance assigned by the prospective mother to accessing as good care as possible, and as early in the pregnancy as possible (Kessner et al., 1973; Chamberlain et al., 1975). On the other hand, since the studies that report a relationship between social factors and neonatal outcomes involved representative populations in which access to medical facilities mirrors that of society in general, it is not surprising that the relationships observed in these studies present the result pattern they do.

In summary, it appears that the generally observed adverse effects of less than optimal social conditions on perinatal outcomes is largely mediated by differences in the quality of the medical facilities available to persons from different social backgrounds rather than by a more direct impact on the mother's environment during pregnancy on the developing fetus. However, it should be emphasized that exceptions to this have been reported in the literature. One such exception is the black subsample of the American Collaborative Project (Niswander & Gordon, 1972). In this sample, the high quality care supplied during pregnancy and delivery served to greatly reduce, but did not completely do away with, the association between adverse social factors and perinatal mortality and neonatal problems. The suggested reason for this is that the environmental conditions influencing the mother through her formative years represented such severe degrees of deprivation or such pervasively aversive life styles, e.g., drug addiction or alcoholism, that she was rendered incapable of producing an optimal baby through an uncomplicated pregnancy, in spite of the best medical care (Kessner et al., 1973).

The Danish data in one instance supported the literature in the area. In one sample, lower birthweight was associated with lower SES. Even if this is not a chance finding, it does not represent a challenge to the notion presented above that SES effects on neonatal outcomes are largely mediated through differential quality of medical treatment. A sizeable literature has shown that in spite of the provision of the

highest quality of medical treatment, smoking during pregnancy, as well as other aspects of the maternal life style, will result in reduced birth-weight (Meyer, Jonas and Tonascia, 1976). These adverse life style charac-teristics are more prominent in lower SES groups (Yerushalmy, 1971). Thus, if any effect of low SES should be expected in the Danish all-white sample, the observed effect on birthweight seems the most probable.

The medical treatment provided to the Danish Perinatal Sample should be expected to have the effect of not only suppressing the observed effect of social variables on the medical outcomes, but also of diminishing the predictability from preceding medical events to medical outcomes measured at a later point in time. This is the logical consequence of instituting preventive and remedial treatment in a sample which is studied over time. For example, since previous reproductive loss (in these analyses represented by the negative data from previous pregnancies variable) is known to be related to increased degrees of complications in later pregnancies (Zachau-Christiansen and Ross, 1975), preventive treatment was instituted in the Danish sample whenever a subjects' medical history included reproductive loss. Each case in which this treatment had an effect serves to reduce the observed relationship between this variable and the pregnancy outcomes analyzed in the present paper. The low predictability of the pre- and neonatal outcomes observed in both of the samples must be interpreted largely to be a function of the syste-matically applied medical treatment procedures. The most striking example of the reduced predictability from earlier to later medical outcome occurrences is reflected by the small amount of the delivery complications variance that was accounted for by the other medical variables, e.g., pregnancy complications. Judging from the literature (Pasamanick, 1954; Towbin, 1978), the predictability from pregnancy complications to delivery complications is considerably diminished in the two samples presented here as compared with the samples usually reported on, i.e., samples not exposed to systematic treatment interventions.

One year status. The results of the analyses of the three one year outcomes are more consistent with the literature than was the case at the neonatal level. It is logical that this would be the case since the conditions of the sample during the first year of life resemble those characteristics of infants in the general population to a much greater extent. That is, after discharge, the mother-child pairs included in the study were left alone by the project staff until the one year follow-up. As expected, the data show that the variables included in this analysis account for a limited proportion of the variance of the one year measures. Consistent with the literature cited previously, one year neurological status shows the lowest degree of predictability (see Table 5). The predictability of one year physical health and motor development was similar in the two samples; approximately 16% and 13% respectively. No association was noted between any of the environmental variables and one year motor development. One year physical health is significantly influenced by one of the environmental variables (daycare); exposure

to institutional daycare is related to an increased incidence of health problems during the infant's first year. No other non-medical variable, including SES, shows a significant path coefficient for the one year health score in either sample. The most likely interpretation of the relationship between insitutional daycare and impaired health status may be related to the fact that being in an institution involves expo-sure to more infections. However, since private daycare experience analyzed in previous analyses has shown no relationship with one year physical health (the variable was therefore excluded from this analysis) it could be that a more psychological interpretation may also be war-ranted. The typical private daycare situation in Denmark is characterized by three to four children in one woman's home for care during the working hours. Thus, the essential difference between institutional and private daycare situations is that the former exposed the infant to more infections, since a larger number of children are kept together. In addition, the institutional daycare situation involves a higher child-adult ratio; thus the amount and continuity of individual attention available for each child over a period of time is probably smaller in comparison with private daycare. In addition to the increased exposure to infections, it may be that the lower degree of personalization of care and inter-action contributes to the poorer general health status of children. Werner, et al., (1971) provide some support for this notion in that their studies showed that family stability constitutes a significant positive influence on child health at age 20 months.

In summary, the analyses of the one year medical outcomes do not present evidence for the notion that environmental influences should be noticeable at an earlier point in time if qualitatively continuous medical outcomes are being analyzed. The significant negative impact of institutional daycare on the physical health of children, however, suggests that further analyses of the effective agents in this relation-ship may produce some support for the early effects of environmental variables on infant health.

Discussion of the Application of Path Analysis

Path analysis has been used in this study in an effort to develop causal models in a substantive field in which very few models exist. Those models which do exist (e.g., the Anderson-Brooks discussion of infant mortality in Sechrest, et al., 1979) are often not specified at the individual level, but at the institutional or state level. Appro-priate information is needed at the individual level for inferences regarding effective intervention to be meaningful. As a multivariate data analytic technique path analysis has many attractive features, but there are several important points that should be considered when using path analysis in an exploratory mode to develop causal models. These cautions relate to the following topics.

Dependency on Previous Data Analyses

In the present study a series of analyses within the inquiry itself was used to help define a general model which could be tested by path analysis. The data analytic work that helped to define the study variables had a profound influence on the final causal modeling that was done. This work, much of which would not be acceptable in a strictly confirmatory mode, gives rise to the distinction between using path analysis for exploratory or for confirmatory purposes.

The preliminary analyses provided the researchers with a clearer understanding of most of the interrelationships among the study variables. In some ways, one of the most useful functions of path analysis has been to provide a distillation of the model. In an exploratory mode, path analysis can be used to define more precisely a proposed causal model and subsequently to modify the model until it provides an accurate description of the data. Cook and Campbell (1971) allude to this procedure. Bentler (1980) and Kenny (1979) provide some guidance in model modification; but the literature has largely ignored the exploratory capabilities of current path analysis procedures. It may be that many published path analysis studies are more exploratory than is readily apparent from the final models that are presented.

Distinguishing Exploratory and Confirmatory Models

The temptation to confuse exploratory and confirmatory models should be resisted. The model presented in this paper is exploratory and should be considered a first step toward a causal model in this field.

Two factors clearly identify the exploratory nature of this model building study. The first is the fact that in general the model was developed on the basis of statistical analyses within the inquiry itself. The second factor is that the initial general model was modified, based on the results of previous tests, to improve the match between the model and the data. Although the total number of parameters changed was relatively small, the model was nonetheless modified; thus, it cannot be considered confirmatory. In an exploratory context the data analysis strategies and outcomes can be very different than for confirmatory analyses. While a certain level of "data snooping" is necessary and desirable in exploratory model building, the final model that is developed can be presented only tentatively. Other researchers must be encouraged to test the model in various ways, e.g., testing the same model on alternative populations or defined subsets, and redefining variables and constructs. Once a model has been proposed in a field, researchers can examine the model in toto or analyze smaller models within the general model, depending upon their particular research interests. Exploratory model building must, by necessity, precede confirmatory efforts in many areas of social science and educationl research.

Defining Appropriate Endogenous Variables

The present study resulted in a model that consists of observed measures. While some effort has been made to construct latent variables, the usefulness of this activity for the one year medical outcomes came into question. While under many circumstances the identification of latent variables or factors is desirable, the present study utilized dependent variables that were essentially the variables of clinical interest. The medical complication scores were derived from information about both the frequency and severity of study variables (in a sense the latent variable was constructed prior to utilizing path analysis). However, causal modeling for one year infant outcomes in the Danish cohort study is but our first step. Further analyses will involve causal modeling for an array of longer term outcomes over a period of 19 years for both mother and child. While the measures of the endogenous varibles are clinically valid and adequate for the one year model, respecification of the measures empirically may provide additional reliability and sensitivity when they are used in models as variables predicting toward longer term child and mother outcomes. Analyses prior to making the causal model explicit will help define the factor and structural sets that will constitute the observed measures defining the latent variables. Such a procedure is in conformance with suggestions by Bentler (1980) and Kenny (1979) in applying path analysis.

Identifying the "Operators" in Exogenous Variables

In the application described here the outcomes of interest were principally medical and physiological over a relatively short temporal span. However, as indicated above, the remainder of our project involves an array of complex outcome variables over a 19 year period for both mother and child. The functions and operations of most of the social variables to be included in this part of the study are complex; so, too, are their causal implications at the individual level of analysis. Thus, analytic attention will be devoted to identifying the specific structures (operators) within a variable that implement a causal relationship prior to making components of the causal model explicit. Analytic dissection of the operators and causal relations is an important precursor to path analysis (Heise, 1975).

Special analytic attention will be given to three problems that we have identified from the analyses described earlier in this paper as having particular importance in making causal inferences over a 19 year period, across many and differing en route variables. The problems are mentioned here simply as additional instruction to the reader.

Identification of the differential functions of a variable over time. For example, mother's age relates strongly to pre- and perinatal

conditions and at the same time it covaries with the psychosocial attri-
butes of the mother that influence post-natal development of the infant.
Thus, very likely, mother's age will have to be treated as both a signal
of biologically related pregnancy complications and as an indication
of the potential influence that social and parenting experiences have
on child and development.

Determination of the relevant attributes of variables that are
global. Variables like mother's age appear to function differently
during the different temporal stages. However, other social variables
are so global that different aspects of the variable may be responsible
for the correlations with the different outcomes over time. For example,
socio-economic status (SES) is a single score scale and is based primarily
on occupational considerations. However, it is conceivable that the
attributes subsumed under SES, e.g., neighborhood, level of education,
appearance of the home, value predispositions of the parent, etc.,
differentially impact on outcome variables at different temporal junctions.
The importance of breaking the global variables down into subordinate
structures as a precursor to causal modeling is clear; so, too, is the
necessity to ensure that the data base will provide access to such specific
elements.

Identification of variables that are moderators of causal relation-
ships. We have called this "sensitization" as a special form of inter-
action. While hierarchical regression analysis appears adequate for
identifying interactions amongst critical variables, "sensitization"
is a special form of interaction requiring ad hoc analyses to ferret
out causal associations. For example, let us say that the variable
"maternal smoking during pregnancy" appears reasonably independent of
the vulnerability to anemia and reflects no statistically significant
relationship with longer term infant health. However, the presence of
maternal smoking coincidental with anemia sensitizes the condition
sufficiently to produce heightened vulnerability now and later on. In
effect, direct attention to this type of problem serves to identify
constellations of variables which define critical subpopulations that
might otherwise get lost in the various residual and error terms.
More importantly for causal modeling, it helps to define new operator
variables that are combinations of aspects of two (or even more)
variables for inclusion in the model at a specified temporal point.

Cross-Validation

Cross-validation strategies should be designed to maximize their
usefulness to both the researchers and their audience. In an explora-
tory mode the validation strategy of splitting a sample or drawing
several random samples provides only limited new information to the
researcher. More frequently, model modification procedures tend to

highlight random differences between samples. While these random dif-
ferences might have heuristic value later on when using model confir-
mation procedures, another strategy in the exploratory mode might be
to define subpopulations of interest on which the model can be tested
and modified. While this strategy does depart from a straightforward
cross-validation, the researcher will obtain more useful information.
Differences that occur between the population and the cross-validation
subsample may reflect random fluctuations or real differences, and can
provide a useful source for generating additional hypotheses. This
procedure would sidestep the longstanding criticism of cross-validation
that it should not be done on the researcher's original population.

Study Type and Data Treatment

Since the purpose of this paper is to describe an application of
path analysis techniques to a prospective longitudinal data base, one
final point should be raised. The Danish cohort data base constitutes
a rich resource of information collected over a lengthy period of time.
While longitudinal in nature, the data collection is, in the main, ex
post facto. The techniques used here are sufficiently robust to provide
only minimum distortions. However, the actual application of path analysis
techniques formally requires interval data. The Danish data base does
not meet that requirement in all cases; it was necessary to treat some
ordinal data items as interval. Asher (1976) and Wilson (1971) indicate
that such treatment will yield findings that are inconclusive since
some legitimate transformations of the ordinal observations will reverse
the conclusions reached. On the other hand, in single equation regres-
sion contexts, Bohrnstedt and Carter (1971) and Labovitz (1967, 1970)
advise going ahead with model testing and parameter estimates. Their
findings indicate that ordinal data can be analyzed by techniques
that formally require interval data without any serious distortion.

The above, coupled with the fact that we are using path analysis
as an exploratory rather than confirmatory technique, might suggest
that other less stringent causal modeling procedures be investigated.
Procedures employing log-linear or logit-linear models might be con-
sidered (Haberman, 1978; and Bock, 1972). These procedures permit
analyses of multidimensional contingency tables, and thus do not
formally require interval data.

Table 1

Variables Included in Analyses of Neonatal Outcomes for the Danish Perinatal Sample

Antecedent Variables

Social data on the family

SES of family at birth of child
Mother's marital status
Mother's age at first birth
Mother's age at birth of
  index child
Father's age at birth of index
  child

Family structure data

Number of older siblings
Spacing of older siblings

Wantedness of pregnancy

Medical factors

Negative outcomes of previous
  pregnancies
Mother's pre-pregnancy health
Parity number
Mental illness in mother

Mediating Variables

Pregnancy complications

Delivery complications

Outcome Variables

Neonatal physical health
  status

Neonatal neurological
  status

Birthweight

187

Table 2

Variables Included in Analyses of One Year Outcomes for the Danish Perinatal Sample

Antecedent Variables

Social data on the family

SES of family at birth of child
Mother's marital status
Mother's age at first birth
Mother's age at birth of index
  child
Father's age at birth of index
  child

Family structure data

Number of older siblings
Spacing of older siblings

Wantedness of pregnancy

Medical factors

Outcomes of previous pregnancies
Mother's pre-pregnancy health
Pregnancy complications
Delivery complications
Birthweight
Multiple birth
Neonatal physical exam score
Neonatal neurological exam score

Mediating Variables

Sociological factors

SES of family at child's one
  year birthday
Child's home environment
Mother's work
Child's exposure to day care

Parental health

Mother's post-partum physical
  condition
Mother's mental illness

Outcome Variables

One year physical health
  status
One year neurological
  functioning
One year motor development
  functioning

Table 3

Glossary of Symbols of Variables

1. Nexib - Spacing to next oldest sibling

2. Moage - Mother's age at birth of index child

3. Prev Health - Mother's pre-pregnancy health problems (scale 0-4 rating severity of illnesses)

4. (-) data - Negative data from previous pregnancies (amount of previous reproductive loss)

5. SES - Socioeconomic status (measured at one year), low score = Hi SES (British Classification of Occupations, 1950)

6. Daycare - Institutional daycare during first year

7. Mo Emp - Mother employed during first year

8. Famsize - Number of older siblings alive at birth of index child

9. Wantedness - Whether or not child was wanted at first hospital visit (higher score = less wanted)

10. Urba/Rurl - Urban vs. rural home address of infant (higher score = rural)

11. PC - a weighted score for pregnancy complications (higher score = more complications)

12. MB - Multiple birth (twin or triplet)

13. DC - a weighted score for delivery complications (higher score = more complications)

14. BW - Birthweight

15. NN - a weighted score for neonatal neurological complications (higher score = more complications)

16. NP - a weighted score for neonatal physical complications (higher score = more complications)

17. 1 yr. Fizz - a weighted score for one year physical health status (higher score = more complications)

18. 1 yr. Neuro - a weighted score for one year neurological status (higher score = more complications)

19. 1 yr. Motor - a weighted score for one year motor developmental status (higher score = more complications)

Table 4
Correlation Matrices for All Variables for Both Samples[1]

	PC	MB	DC	BW	NN	NP	1yr Fizz	1yr Neur	1yr Mot	Nex Sib	Mo Age	Pre Ill	Pre Preg	SES	Day care	Mo Emp	FS	Want	U/R
Pregnancy Complications (PC)	--	12	18	00	-06	07	02	03	04	30	27	07	08	-12	-02	-03	00	-13	02
Multiple Birth (MB)	04	--	04	-28	06	07	05	-04	14	02	03	-06	-02	03	-05	-03	07	-03	-06
Delivery Complications (DC)	19	14	--	-10	19	38	17	23	21	12	13	13	06	-03	03	-04	-03	-09	-08
Birthweight (BW)	-03	-20	-07	--	-32	-30	-19	-02	-25	09	12	00	-03	-10	00	10	02	03	03
Neo-Neurological (NN)	05	14	23	-32	--	53	18	20	17	-05	-04	07	-02	05	-01	06	-03	-11	-05
Neo-Physical (NP)	09	11	37	-29	51	--	27	20	26	-05	01	08	-08	02	00	-02	13	-07	05
One Year Physical (1yr Fizz)	04	02	03	-21	22	10	--	33	33	-10	-13	00	-05	11	23	14	-03	07	-09
One Year Neurological (1yr Neur)	04	-03	05	08	11	13	19	--	29	-05	-03	04	-06	-03	07	06	01	06	-02
One Year Motor (1yr Mot)	03	12	10	-25	23	25	19	06	--	-12	-05	-06	-05	-12	03	00	-05	02	-06
Spacing to Next Child (Nexsib)	21	00	07	03	-05	10	07	05	04	--	58	05	13	-05	06	05	-05	01	-06
Mother's Age (Moage)	19	01	18	12	-02	13	-06	12	05	53	--	11	29	-32	00	-11	30	-16	-07
Previous Illness (Preill)	09	-06	21	04	11	08	09	00	09	11	12	--	10	01	-05	-07	-04	-14	-04
Previous Pregnancy Problems (Prepreg)	09	07	07	-08	06	07	10	07	15	27	34	03	--	-03	-02	-11	14	01	-10
One Year SES (SES)	-05	05	-08	-20	02	04	16	01	03	-09	-36	-07	-01	--	02	18	09	33	-06
Daycare (Daycare)	03	-07	03	-10	-01	-04	29	-07	-07	-11	-11	-07	04	15	--	37	-09	10	-08
Mother's Employment (Mo Emp)	-05	-03	-03	-07	-08	-05	21	-07	-06	-08	-20	-07	01	26	49	--	-01	18	-06
Family Size (FS)	15	07	06	-07	03	19	04	-01	10	55	46	02	32	-01	-13	-14	--	33	-07
Wantedness of Infant (Want)	-06	03	-12	-03	-07	-07	07	-04	-02	-01	-26	-08	-05	38	15	12	16	--	-04
Urban-Rural (U/R)	04	03	09	05	01	08	-09	-01	05	00	-05	06	-05	-11	-13	-16	05	02	--

[1] Sample 1 correlations are above the diagonal;

Table 5

Multiple R^2 for Dependent Variables

Outcome Variable	Variance Explained (R^2)	
	Sample 1	Sample 2
Pregnancy Complications	.12	.05
Delivery Complications	.06	.10
Birthweight	.09	.08
Neonatal Neurological	.14	.14
Neonatal Physical	.26	.23
One Year Physical	.16	.17
One Year Motor	.13	.11
One Year Neurological	.08	.02

Figure 1 - Time Frame for Path Diagrams

X Variables (with unknown or variable temporal characteristics)	Y Variables (with known temporal sequencing)			
	Time 1 Prior to Birth	Time 2 At Birth	Time 3 After Birth	Time 4 One Year
Mother's age Previous health (-) data previous pregnancies SES Day Care Home environment Mother's employment Family size Wantedness Urban/rural	PC MB	DC BW	NN NP	1 yr (outcome)
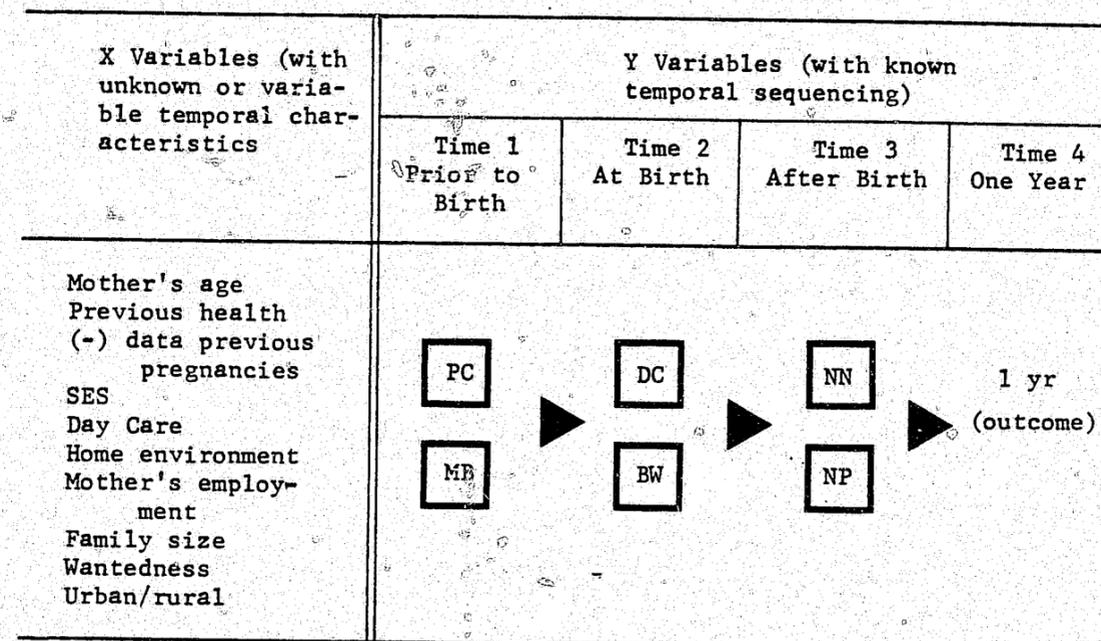
Figure 2

Discriminant Function Analyses: A Temporal Composite of

Seven Medical Outcome Variables
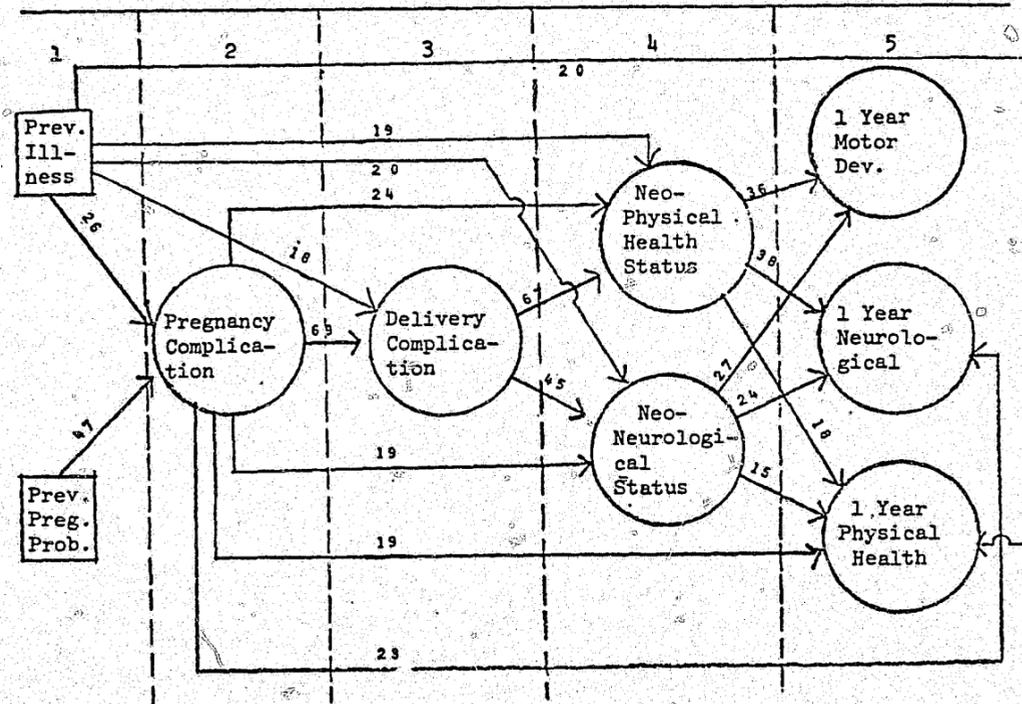
Predictor Variable	PC	DC	NP	NN	FIZZ	NEGRO	MOTOR
Multiple Birth	-55	-16	D	-05	-04	-07	-11
No. of Sibs	11	42	03	25	-33	16	10
Mother's Age	-46	-49	D	06	19	-20	-13
Birth Weight	15	12	66	77	38	D	53
SES First Yr.	--	--	--	--	-25	16	15
Parity Number	D	D	-16	-31	D	13	-22
Wantedness	--	--	--	--	-17	D	08
Sex of Index	D	D	12	D	27	31	D
Mother's Emploc.	--	--	--	--	-17	10	06
Day Care Inst.	--	--	--	--	-45	D	D
SES Bio-Father	09	-11	06	D	-09	D	D
Prev. Preg. Prob.	-47	D	D	D	-05	06	D
Prev. Health	-26	-18	-19	-20	-20	D	D
Preg. Comp.	--	-69	-24	-19	-19	-23	D
Delivery Comp.	--	--	-67	-45	D	D	D
Neo-Physical	--	--	--	--	-18	-38	-36
Neo-Neuro.	--	--	--	--	-15	-24	-27
NexSib	28	23	D	D	D	D	D

Notes: (1) -- = variable not included in the analysis
(2) D = dropped from analysis due to non contribution
(3) Numbers are canonical (standardized) DF coefficients

Figure 3- Initial Generalized Model
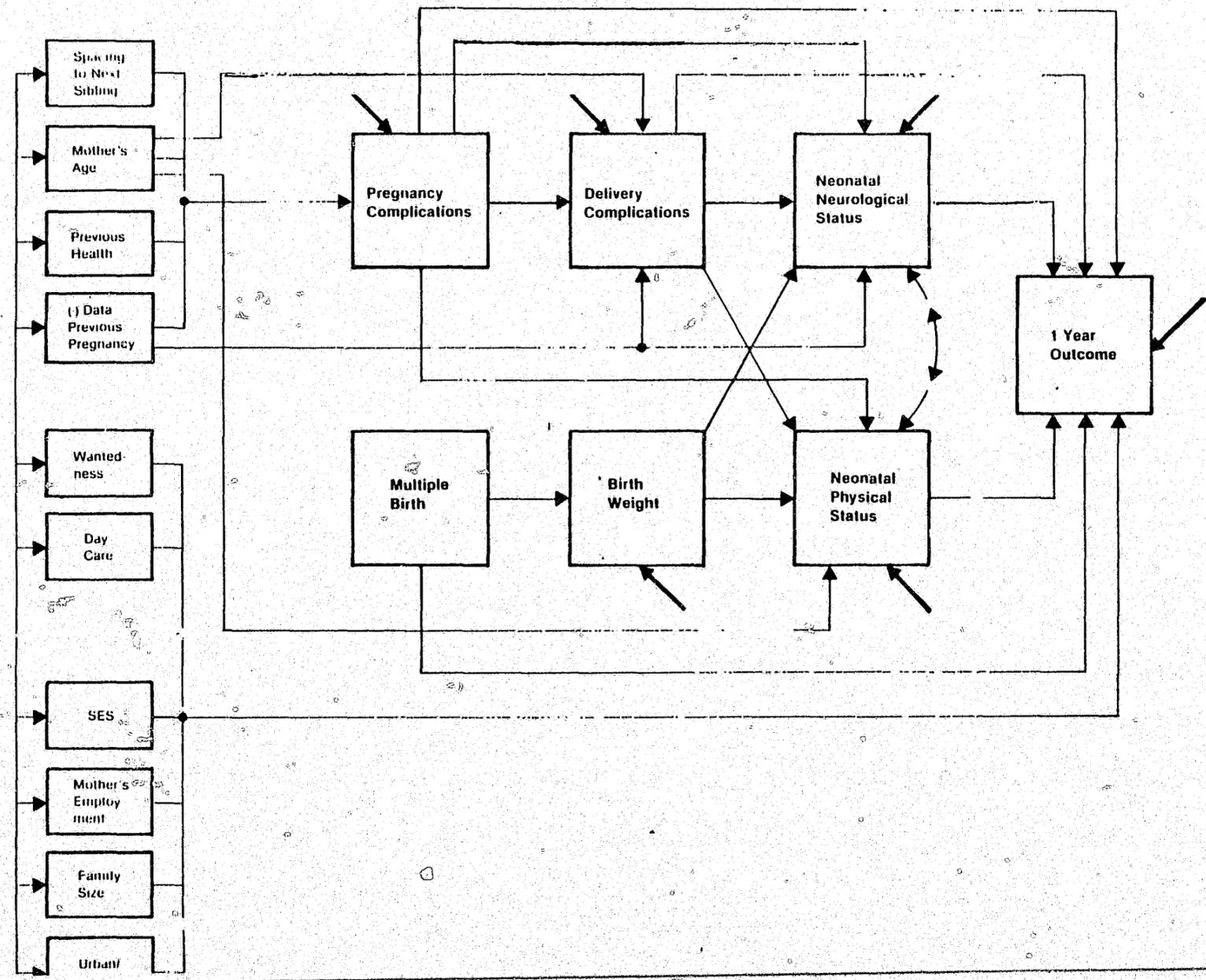
Figure 4 - Model Predicting to One Year Outcomes: Sample 1
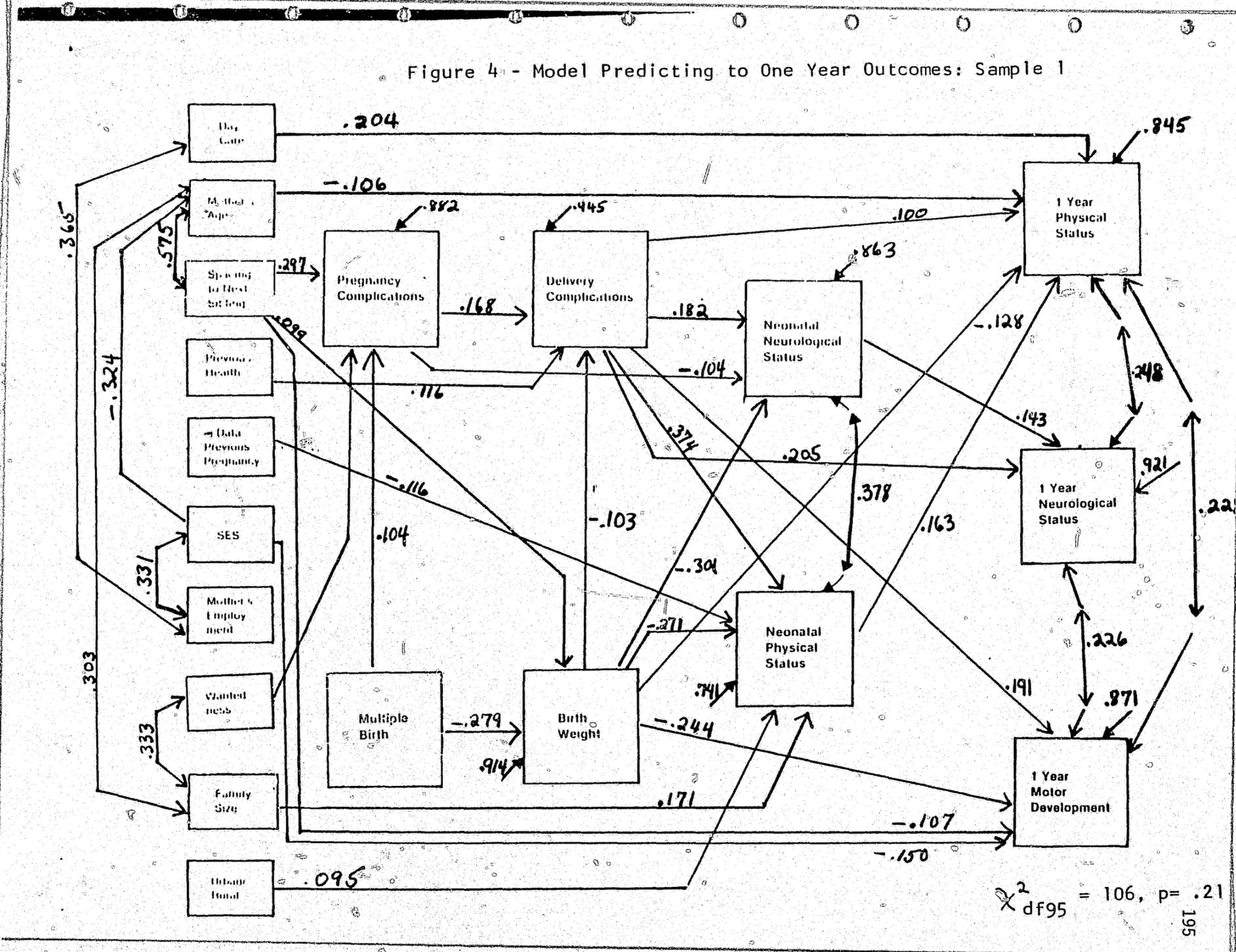
Figure 5 - Model Predicting to One Year Outcomes: Sample 2

$\chi^2_{df102} = 117$, p= .15

196
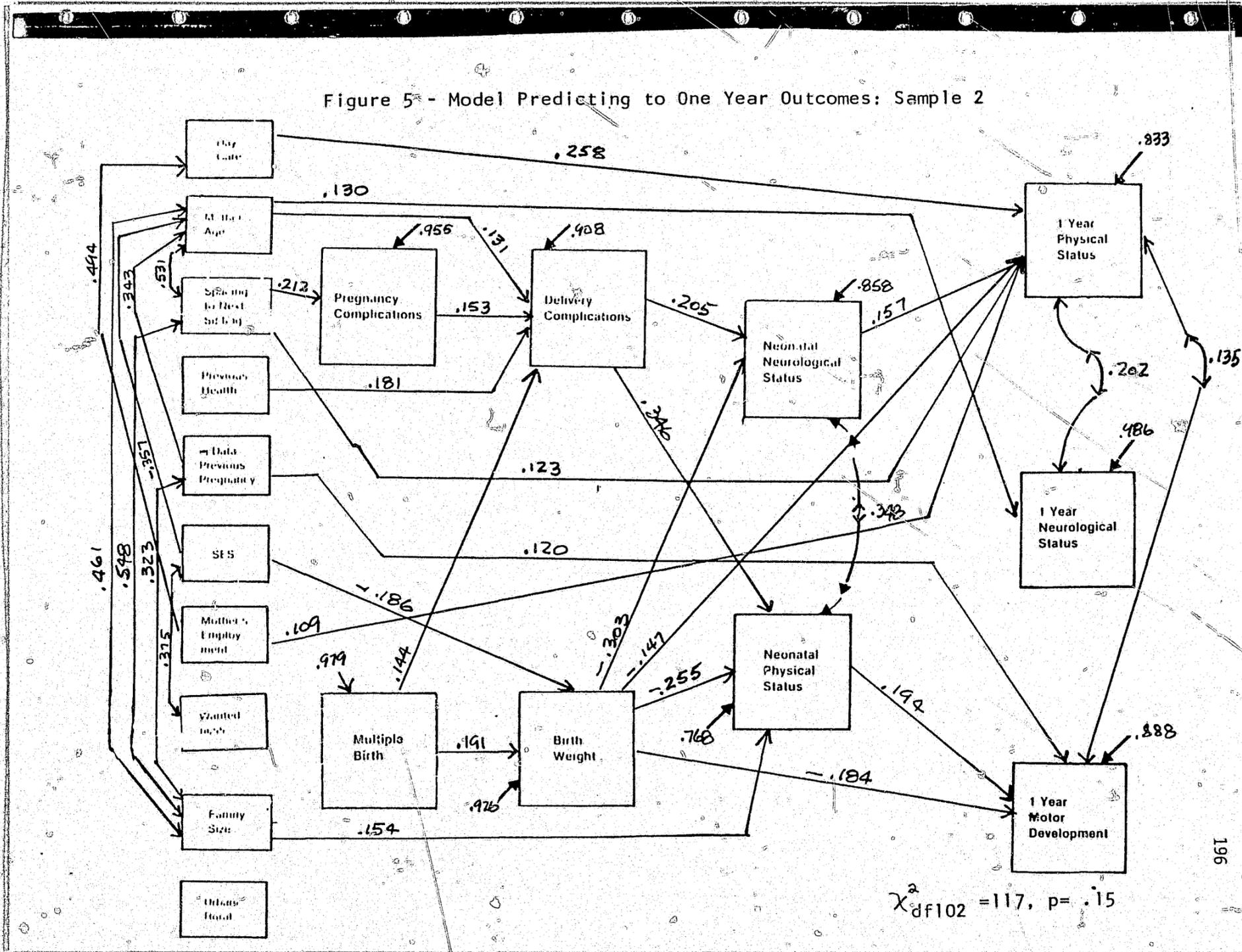
Figure 6 - Model Showing Paths Significant in Both Samples
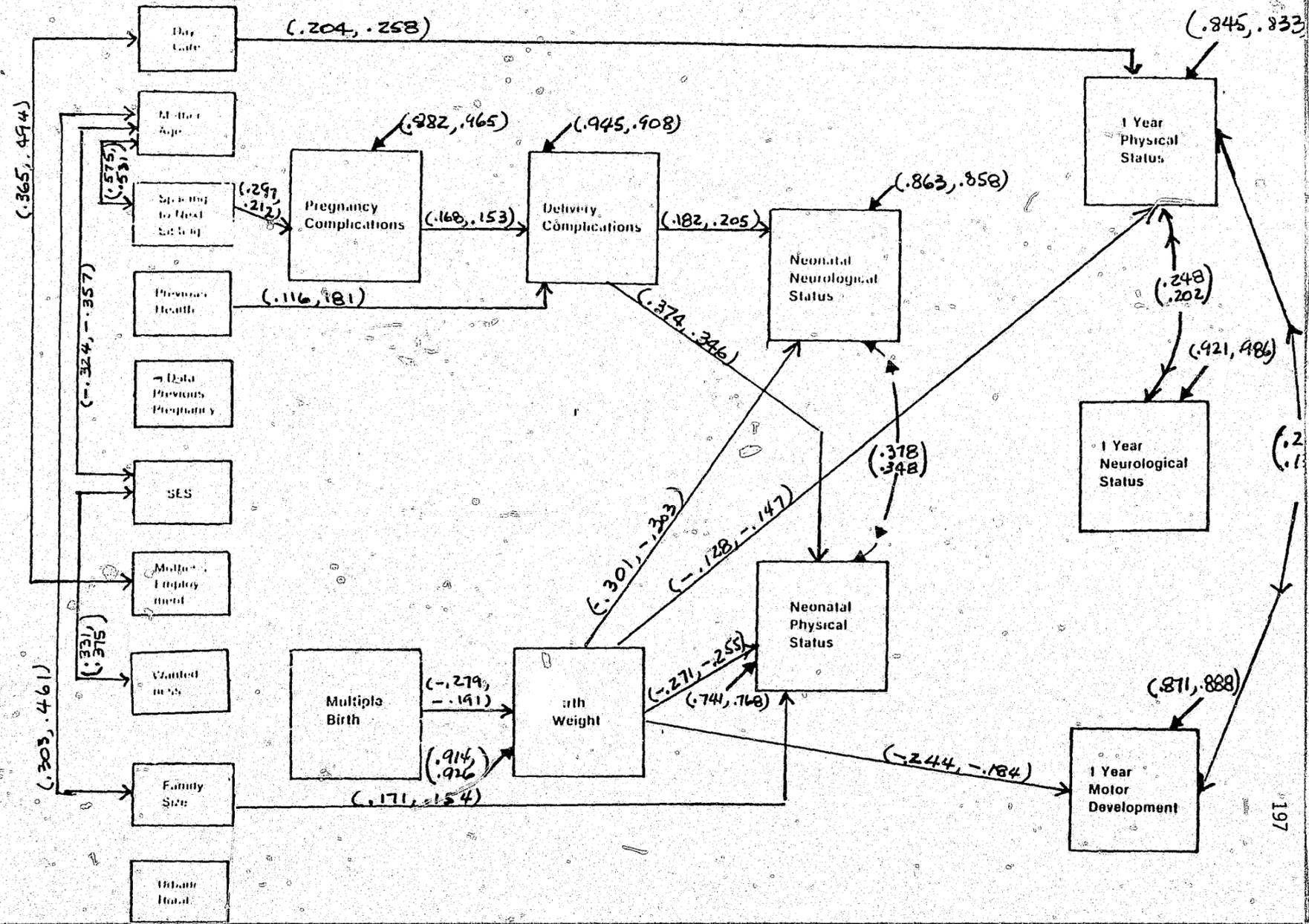
References

Asher, H. B. Causal Modeling. Beverly Hills: Sage, 1976.

Bentler, P. M. Multivariate analysis with latent variables: causal modeling. Annual Review of Psychology, 1980, 31.

Bock, R. D. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. Psychometrika, 1972, 37, 29-51.

Bohrnstedt, G. W. & Carter, T. M. Robustness in regression analysis. In H. L. Costner (Ed.), Sociological methodology. San Francisco: Jossey-Bass, 1971.

Bradley, H. R., Caldwell, B. M. & Elardo, R. Home environment and cognitive development in the first 2 years: a cross-lagged panel analysis. Developmental Psychology, 1979, 15(3), 246-250.

Bronfenbrenner, U. A report on longitudinal evaluations of pre-school programs. DHEW Publication No. (OHD) 75-25, 1974.

Butler, N. R. & Alberman, E. D. (Eds.) Perinatal problems. London: Livingstone, 1969.

Chamberlain, R., Chamberlain, G., Howlett, B. & Claireaux, A. British births, 1970. London: Heinemann, 1975.

Cohen, J. & Cohen, P. Applied multiple regression: correlation analysis for the behavioral sciences. New York: Wiley, 1975.

Cook, T. D. & Campbell, D. T. Quasi-experimentation: design and analysis issues for field settings. Chicago: Rand McNally, 1979.

Crano, W. What do infant mental tests test? A cross-lagged panel analysis of selected data from the Berkeley growth study. Child Development, 1977, 48, 144-151.

Deegan, J. The effects of multicollinearity and specification error on models of political behavior. Ph.D. dissertation, Univ. of Michigan, 1972.

Dijkstra, J. De prognostische betekenis van neurologische afwijkingen bij pasgeboren kindern. V. R. B. Kleine Der A 3-4 Gronigen, 1960.

Drage, J. S. & Berendes, H. Apgar scores and outcome of the newborn. Pediatrics Clinics of North America, 1966, 13, 635-643.

Field, T., Hallock, N., Ting, G., Dempsey, J., Dabiri, C. & Schuman, H. H. A first-year follow-up of high-risk infants: formulating a cumulative risk index. Child Development, 1978, 49, 119-131.

Fitzhardinge, P. M. & Ramsey, M. The improving outlook for the small prematurely born infant. Developmental Medical Child Neurology, 1973, 15, 447-459.

Goldstein, K. M., Caputo, D. V. & Taub, H. B. The effects of prenatal and perinatal complications on development at one year of age. Child Development, 1976, 47, 613-621.

Haberman, S. J. Analysis of qualitative data: Vol. 1, introductory topics. New York: Academic Press, 1978.

Heber, R. Intensive intervention programs prevent retardation. APA Monitor, 1975, 7(9 & 10).

Heise, D. R. Causal analysis. New York: Wiley, 1975.

Honzik, M. P., Hutchings, J. J. & Burnip, S. R. Birth record assessments and test performance at eight months. American Journal of Diseases of Children, 1965, 109, 416-426.

Joreskog, K. G. & Sorbom, D. LISREL IV: Analysis of linear structural relationships by the method of maximum likelihood. Chicago: National Educational Resources, 1978.

Kenny, D. A. Correlation and causality. New York: Wiley, 1979.

Kerlinger, F. N. & Pedhazur, E. J. Multiple regression in behavioral research. New York: Holt, Rinehart and Winston, 1973.

Kessner, D. M., Singer, J., Kalk, C. E. & Schlesinger, E. R. Infant death: an analysis by maternal risk and health care. Washington, D. C.: National Academy of Sciences, 1973.

Labovitz, S. The assignment of numbers to rank order categories. American Sociology Review, 1970, 35, 515-524.

Labovitz, S. Some observations on measurement and statistics. Social Forces, 1967, 46, 151-160.

McGarvey, W. E. Longitudinal factors in school desegration. Unpublished doctoral thesis, Univ. of Southern California, 1977.

Mednick, B. Intellectual and behavioral functioning of 10-11 year old children who showed certain transient neurological symptoms in the neonatal. In S. Chess and A. Thomas (Eds.), Annual progress in child psychiatry and child development. New York: Brunner/Mazel, 1978, 59-76.

Mednick, B. Intellectual and behavioral functioning of ten- to twelve-year old children who showed certain transient symptoms in the neonatal period. Child Development, 1977, 48, 844-853.

Mednick, B., Baker, R. L. & Sutton-Smith, B. Teenage pregnancy and perinatal mortality. Journal of Youth and Adolescence, 1979, 8, 343-357.

Mednick, B. & Michaelsen, N. Neurological and motor functioning of 10-11 year old children who showed certain transient neurological symptoms in the first five days of life. Acta Neurologica Scandinavica, 1977, 56, 70-78.

Mednick, S. A., Mura, E., Schulsinger, F. & Mednick, B. Perinatal conditions and infant development in children with schizophrenic parents. Social Biology, 1971, 18, Supplement, 103-113.

Meyer, M. B., Jonas, B. S. & Tonascia, J. A. Perinatal events associated with maternal smoking during pregnancy. American Journal of Epidemiology, 1976, 103(5), 464-476.

Niswander, K. R. & Gordon, M. The women and their pregnancies. Philadelphia: Saunders, 1972.

Niswander, K. R., Friedman, E. A., Hoover, D. B., Pietrowski, R. & Westphal, M. Fetal morbidity following potentially anoxigenic obstetric conditions. I. Abrupto placentae. II. Placenta previa. III. Prolapse of the umbilical cord. American Journal of Obstetrics and Gynecology, 1966, 95, 838-846.

Parmelee, A., Kopp, C. & Sigman, M. Selection of developmental assessment techniques for infants at risk. Merrill-Palmer Quarterly, 1976, 22, 177-199.

Parmelee, A. & Michaelis, R. Neurological examination of the newborn. In J. Hellmuth (Ed.), Exceptional infant. Vol. 2. New York: Brunner/Mazel, 1971.

Pasamanick, B. Epidemiology of behavior disorders. Neurology and psychiatry in childhood. New York: Williams & Williams, 1954, 397-403.

Pharoah, P. O. D. Obstetric and neonatal care related to outcome. A comparison of two maternity hospitals. British Journal of Preventive Social Medicine, 1976, 30, 257-261.

Prechtl, H. Problems of behavioral studies in the newborn infants. In D. Lehrman, R. Hinde & E. Shaw (Eds.), Advances in the study of behavior. Vol. 1. New York: Academic Press, 1965 (a).

Rubin, R. A., Rosenblatt, C. & Balow, B. Psychological and educational sequelae of prematurity. Pediatrics, 1973, 52(3), 352-363.

Sameroff, A. J. The etiology of cognitive competence: a systems perspective. In R. B. Kearsley & I. E. Sigel (Eds.), Infants at risk. New York: Wiley, 1979.

Sameroff, A. J. Infant risk factors in developmental deviancy. In E. J. Anthony & C. Koupernik (Eds.), The child in his family (Vol. 4). New York: Wiley, 1978.

Scarr-Salapatek, S. & Williams, M. L. The effects of early stimulation on low birthweight infants. Child Development, 1973, 44, 94-101.

Sechrest, L., West, S. G., Phillips, M. A., Redner, R. & Yeaton, W. (Eds.), Evaluation studies review annual, Vol. 4. Beverly Hills: Sage, 1979, 238-283.

Sorbom, D. Detection of correlated errors in longitudinal data. British Journal of Mathematics and Statistical Psychology, 1975, 28, 138-151.

St. Claire, K. Neonatal assessment procedures: a historical review. Child Development, 1978, 49, 280-292.

Towbin, A. Cerebral dysfunctions related to perinatal organic damage: clinical-neuropathologic correlations. Journal of Abnormal Psychology, 1978, 87(6), 617-635.

Werner, E. E., Biermen, J. M. & French, F. E. The children of Kauai. Honolulu: University of Hawaii Press, 1971.

Wilson, T. P. Critique of ordinal variables. In Blalock, H. M. (Ed.), Causal models in the social sciences. Chicago: Aldine-Atherton, 1971.

Wright, S. The method of path coefficients. Annual Mathematical Statistics, 1934, 5, 161-215.

Yerushalmy, J. The relationship of parents' cigarette smoking to outcome of pregnancy--implications as to the problem of inferring causation from observed associations. American Journal of Epidemiology, 1971, 93(6), 443-456.

Zachau-Christiansen, B. & Ross, E. M. Babies: human development during the first year. New York: Wiley, 1975.

Chapter 8

Some Practical Suggestions for

Minimizing Subject Attrition

Michele Harway

Social Science Research Institute

University of Southern California

Michele Harway

Center for Longitudinal Research

University of Southern California

Most longitudinal research projects involve relatively long periods of observations of their samples with measurements taken at at least two points in time. Long periods of observation often lead to considerable attrition of subjects and make attrition a major problem of longitudinal research. Two of the reasons attrition occurs are: 1. Respondents may refuse continued participation in the project, perhaps because they get tired of repeated questioning; 2. Cohort members move, change names, go to jail or mental hospital or die. In our highly mobile society, it is likely that moving will be the cause of a large portion of the attrition in any one project. Because of the difficulties inherent in tracking mobile cohort members, and because of the expense in reaching them for questioning in their new locations, they are often dropped from the cohort.

Since certain types of individuals are more likely to refuse continued participation in a longitudinal project, and the disadvantaged and single-parent families are more likely to be difficult to locate, attrition from the sample is not a randomly distributed event. Thus, generalizing from a cohort where attrition of a considerable number of cases has occurred is likely to result in errors of interpretation. Such errors are particularly problematic in pursuing certain types of research questions; less problematic for others. Thus, if we are interested in the incidence of delinquency or mental illness in a long-term longitudinal sample (as representative of a larger population), attrition becomes particularly problematic. Our measure of incidence will be biased to the extent that attrition has not been a random event. If we are studying delinquency, it is especially likely that individuals who are at the greatest risk for becoming delinquents (those who are from poor or single-family homes) will also be those who drop out.

In spite of attrition, we may be able to study problems that are appropriate to the composition of the remaining cohort. Thus, if intact middle-class families tend to remain in the sample, we could study antecedents of delinquency which are specific to those families.

Because attrition is perceived as such a critical problem in longitudinal research, much care and attention is usually devoted to how cohort members are handled at the initial contact and how they will be followed-up through time so as to minimize later attrition. Longitudinal researchers have developed a variety of techniques designed to minimize subject loss. Some of these techniques are described below.

In order to identify some of the ways that researchers in the field have minimized subject loss, we conducted an informal survey among the longitudinal researchers with whom we had been corresponding. The survey explored five areas relating to subject identification, retention and follow-up. Specifically, respondents were asked to describe ways of:

1. Identifying all members of a population
2. Locating these individuals
3. Convincing subjects to participate in the study
4. Laying the groundwork at the initial contact for successful follow-up
5. Ensuring participation at follow-up

The suggestions which follow come from one or more of the 74 individuals who returned our questionnaires. While no information is available regarding empirical tests for most of these suggestions, they are presented here in the hope that they will benefit future longitudinal researchers.

Identifying Members of a Population

Most suggestions for methods of identifying members of a population revolved around the use of existing records or registers which form part of the public domain. While many of these registers are not representative, they can be useful for specific purposes. These registers include:

1. Telephone books or city directories
2. Lists from the census
3. Postal rosters
4. School rosters
5. Organization membership lists
6. Registers of births
7. Marriage records
8. Other vital statistics records
9. Medical records, hospital and psychiatric records

Other recommendations involve contacting key informants who could identify members of a population of interest: for example, obtaining lists of pregnant women from their physicians, or lists of members of an organization from key leadership.

Techniques which were likely to reach a haphazard collection of individuals were also described. These involved attaching message cards to door knobs, knocking on all doors in a chosen community, mass mailings, radio messages, or identifying some event which members of the population would attend and contacting them there. Finally, talking to historical and geneological societies whose members are familiar with different ways of locating and identifying individuals was proposed as a potentially fruitful source of new techniques to locate population members. Methods such as these do not yield representative samples but may be useful in identifying hard-to-find populations (e.g. identical twins).

Locating Members of a Population

Survey respondents suggested four ways of locating samples:

1. Through records. Some of the types of records which survey respondents found useful in locating members of their cohort included:

a. special directories (e.g., directories of professional organizations);
b. insurance and claim records;
c. draft board and military records;
d. IRS searches, and
e. state driver's license records

2. By contacting agencies. The types of agencies suggested varied a great deal depending on the nature of the project. However, these included:

a. schools
b. churches
c. post offices
d. alumni agencies
e. credit bureaus
f. places of employment
g. corectional institutions
h. the armed services
i. labor unions
j. professional tracing firms (which are said to cost between $5 and $10 per name).

3. By contacting friends, relatives, employers, others. Additional individuals who are listed as potentially of help in locating subjects include:

a. neighbors
b. police chiefs
c. physicians

4. By writing to the subject at a previously known address or telephone number. The mail technique is likely to yield information on any new address if certified mail is used or if mailings are sent bulk mail with address correction requested.

Convincing subjects to participate in the study, setting the stage for successful follow-ups and ensuring continued participation

The most consistent finding of this portion of the survey was the extent to which researchers stress the personal touch in appealing to subjects to participate in their study, both initially and at follow-up.

A typical suggestion for convincing subjects to participate was that forwarded by one survey respondent: "We have found that taking the time to inform the subjects about what you are doing and why, and basically letting them know how much you need their help, is the best method--better than paying them, even. The best possible method is to interest them in what you are doing so they want to be part of it and feel it is relevant to their lives and interests." Another researcher notes that: "... We have spent quite a lot of time listening to people. In some cases, we have had Christmas presents, letters, etc., from people who came to see us as resources for their lives. The essential thing is to recognize that people in longitudinal research, on whom large demands are made, are not 'research subjects' and should not be manipulated." Thus, the personal touch is urged as a method of getting people

to participate in longitudinal studies. Personalizing requests for participation in the study is deemed much more effective a method than appealing to individuals' altruistic motives. One researcher comments that "appeals based on somewhat abstract humanitarian goals (e.g., society can better plan services for the entire population of older people if we can accurately understand the patterns of behavior) do not help as much as personal appeals (e.g., please help me; we are trying to do a very thorough and accurate job in collecting this information, and it would help me tremendously if you would consent to participate.)."

Similarly, the personal nature of the initial contact was stressed as the single most important element to ensure successful follow-ups. Continued personal contact is also viewed as important. One study respondent notes: "We have found that the more contact subjects have with us, the more they are willing to continue participating in follow-ups. I think this is largely because we treat them with respect, appreciate their assistance, arrange things at their convenience and inform them of our results." Birthday and Christmas cards, personal letters of thanks, long distance telephone calls between Christmas and New Year's, regular phone contact every three months are some of the personal techniques mentioned here.

In addition to invoking the personal touch, two other approaches increase subject cooperation. These include providing incentives (listed below) and making it easier for subjects to participate (described in a subsequent paragraph).

The incentives suggested range from monetary payment to the provision of some free service, such as a free medical examination, free counseling or free speech and hearing tests. Some researchers mentioned giving gifts to subjects, providing them with magazine subscriptions, taking them on excursions, inviting them to conferences to share results of the study with them and, of course, providing reports of study results. Finally, one researcher noted that in his project, research staff occasionally do favors for elderly subjects, such as giving them rides, getting medicine or food, etc.

Participation is also encouraged by being accommodating to the subjects' needs. This included working interviews or examinations around subjects' schedules and geographical locations, picking up and delivering subjects, and going to their homes or offering to pay their expenses, even minor ones such as parking.

All of this personal contact and effort to maintain as close to intact a sample as possible is time consuming and expensive. Notes one researcher: "Granting agencies must be sensitive to the costly and time-consuming problems of sample retention and must fund accordingly." Researchers from their end, must be sensitive to the importance of laying the groundwork for follow-ups and the importance of maintaining as complete a sample as possible. In the following sections we will discuss the efficacy of different types of follow-up techniques.

Follow-up techniques

Several longitudinal researchers, aware of the critical importance of efficient follow-up procedures in the maintenance of a scientifically valid sample, have themselves researched the efficacy of follow-up techniques.

In addition to studies which have examined the effects of volunteer bias (Rosenthal & Rosnow, 1969; Cochrane & Duffy, 1974) there are a number of investigators who have looked at the effects of respondent non-participation in longitudinal follow-ups and the efficacy of various follow-up techniques.

Paikin, Jacobsen, Schulsinger, Godtfredsen, Rosenthal, Wender & Kety (1974) examined the characteristics of the group of individuals who refused participation in a longitudinal study of adoption. Because of the particular nature of the data base from which these investigators identified their subject pool--the population registers uniformly maintained on Danish citizens--it was possible to conduct statistical tabulations of characteristics of "refusals." Some of the demographic characteristics which differentiated between participants and refusals included: marital status (more unmarried refused), parental status (non-parents refused more often) and social class (significantly more male refusals were categorized as being in the lower social classes.).

In addition to studying characteristics of refusals, these researchers also studied the tenacity of the refusals in the face of repeated attempts at persuasion to participate. These repeated attempts were conducted over a period of several years. Neither financial incentives nor the personal touch, which was described in earlier sections as so effective in maintaining a sample, were successful in convincing refusals to participate in the study.

What these data suggest is that continued attempts at persuasion to participate in a longitudinal study do not yield an increased participation rate. However, to the extent that these data are generalizable to other studies, it appears that even a substantial initial refusal rate (20% in this study) only impacts specific determinants of the representativeness of the initial sample.

Several studies report on the feasibility of tracing respondents in longitudinal studies. Mitchell and Jackson (1978) tracked a sample of British school children after a fifteen year lapse in contact. They decided to concentrate on the most atypical children in the original sample since "the children with the 'most problems' might also be the most mobile" (p. 510). Five percent of the original sample were included in this study. The authors are optimistic about the feasibility of tracing school-aged children into adult life even with the large time lag reported, since 88% of the children tracked were successfully located.

In the United States, Crider and Willits (1973) report on an attempt at locating subjects more than a decade after they had last been surveyed. Three tracking methods were used: mail, telephone calls and community visits. Success rates were lowest with those sought through a mail track (48% located), while 80% were located with long distance telephone methods and 81% through community visits. Clearly telephoning is likely to be considerably less expensive and time consuming than community visits. When all three techniques were combined, a 97% success rate was obtained. Success rates varied by rurality (with locating rates higher in rural areas) but not generally by individual characteristics of respondents such as sex, marital status, age and occupation.

The results of both studies reported above are quite encouraging. If follow-ups conducted ten to fifteen years after a previous contact yield such high returns, researchers who plan their studies and follow-ups carefully and make frequent intermediate contacts are likely to be able to duplicate these success rates. Moreover, in situations where intermediate contacts are not possible, it seems that it would still be feasible to locate some substantial portion of the original sample.

References

Cochrane, R. & Duffy, J. Psychology and scientific method. <u>Bulletin of the British Psychological Society</u>, 1974, <u>27</u>, 117-121.

Crider, D.M. & Willits, F.K. Respondent retrieval bias in a longitudinal survey. <u>Sociology and Social Research</u>, 1973, <u>58</u>(1), 56-65.

Mitchell, S. & Jackson, D. Preliminary communication: The feasibility of tracing a sample of school boys after a lapse of fifteen years. <u>Psychological Medicine</u>, 1978, <u>8</u>, 509-513.

Paikin, H., Jacobsen, B., Schulsinger, F., Godtfredsen, K., Rosenthal, D., Wender, P. & Kety, S.S. Characteristics of people who refuse to participate in a social and psychopathological study. S.A. Mednick, F. Schulsinger, J. Higgins & B. Bell (Eds.) In <u>Genetics, Environment and Psychopathology</u>, Amsterdam: North-Holland Publishing Company, 1974.

Rosenthal, R. & Rosnow, R.L. The volunteer subject. In R. Rosenthal and R.L. Rosnow (Eds.), <u>Artifacts in behavioral research</u>, New York: Academic Press, 1969.

Chapter 9

Usefulness of Official Records in Longitudinal

Research in Criminology

Margaret Heim

University of California, Los Angeles

and

Katherine S. Teilmann

Social Science Research Institute

University of Southern California

83032

Usefulness of Official Records in Longitudinal

Research in Criminology

Margaret Heim

University of California, Los Angeles

and

Katherine S. Teilmann

Social Science Research Institute

University of Southern California

This book presents a wealth of longitudinal data bases for further analyses. This chapter will explore issues pertaining to the use of these data bases for the study of criminal behavior. These issues will be discussed with the assumption that the reader has little or no background in criminology.

Four areas will be presented:

1) Types of research questions about criminal behavior that may be of interest with a longitudinal data base

2) The relative merits of two types of measures of criminal activity, self-report and official records

3) Available resources of official records in the geographic locations of the data bases presented in this book

4) Problems and shortcomings a researcher would encounter in using specific official records.

Types of Research Questions

Some of the data bases in this book include variables of traditional interest in the study of criminal behavior: psychiatric disorders, personality variables, mental abilities, family structure and practices, socioeconomic status, level of education and ethnicity. These variables may be examined as predictors of criminal behavior. Other studies involve particular interventions that may affect the incidence of criminal behavior. Studies that collect data from different generations of families and those collecting medical and physiological information may allow the examination of the role of biological factors in the incidence of criminal behavior.

Not only can the researcher use these data bases to explore the antecedents of the incidence of criminal behavior, he or she may want to examine the antecedents of specific types of criminal behavior. For example, a researcher may be interested in whether the level of social-emotional anxiety in childhood predicts to later violent behavior. Or, a researcher may examine whether differences in level of intelligence predict differences between those committing a violent crime and those committing a property crime.

Chronicity and patterns of criminal behavior may also be of interest to a researcher. A researcher may want to distinguish between the one-time thief and the habitual thief or between a person committing only one type of crime and a person committing many types of crime.

In summary, at least four types of questions may be of interest about criminal behavior in relation to a longitudinal data base. These questions involve the 1) incidence, 2) type, 3) chronicity and 4) patterns of criminal behavior.

Self-Report Vs. Official Records

One of the first issues a researcher of criminal behavior must consider is how criminal behavior will be operationally defined. Two of the most common operationalizations are individual's self-report of criminal activity and official arrest records. A third measure of criminal activity, the victimization survey, is omitted here because its results cannot be associated with the person committing the crime. However, studies comparing victimization surveys with official records can be useful in determining the validity of official records.

Self-reports and official records as measures of criminal activity are typically seen in opposition to each other with one's benefits as the other's deficits. Since there is a tradeoff in the choice of one over the other, the choice of measure might well be affected by the particular research question. Understanding some of the assets, the problems and sources of bias in each measure should help the researcher make an intelligent choice for a particular research study.

One major criticism leveled against official records is that they reflect only a small proportion of all crimes committed. Loss of information in official records can occur at several stages. First, not all crimes are detected or recognized as crimes. Second, of those detected, not all are reported to the police. Third, of those reported, not all lead to an arrest. Fourth, of those arrested, not all are reported by the police to a central records agency. Fifth, of those reported, not all are acceptable to the records system and thus are not retrievable from the central agency.

Underreporting of crime in official records will affect some studies more than others. For example, the study that wants to maximize the probability of discovering whether individuals in the sample committed a crime would not get the best results from official records, particularly if the sample size was small. Few in a small sample from the general population will have been arrested. However, with a larger sample or a sample selected from a group which tends to be overrepresented in official arrest records, official records may be an adequate measure.

In addition to general loss of information, there may be particular biases at each stage of loss in official records based on the type of crime or the person committing the crime. For example, a visible crime such as disturbing the peace is more likely to be detected and reported than a less visible crime such as obscene phone calls. Witnessed crimes may be more likely to yield an arrest than unwitnessed crimes. A serious crime such as murder is more likely to be reported to a central records bureau than a less serious crime such as shoplifting. The seriousness of the crime has been suggested as one of the most important factors associated with bias in official records (Skogan, 1977; Booth, Johnson & Choldin, 1977). Aspects about the person such as ethnicity, gender and frequency of committing offenses may also be sources of bias in official records (see Empey, 1978).

Several studies (Uniform Crime Reports, 1976; Wolfgang, Sellin & Figlio, 1972; Hirschi, 1969) have shown that non-whites, particularly blacks, appear disproportionately on arrest records in that the proportion of blacks arrested is much higher than the proportion of whites arrested. In studies using self-report among adolescents, the reports of crime become less disproportionate with respect to ethnicity. Blacks report only a slightly higher frequency of crime than do whites (Gold, 1970; Hirschi, 1969; Williams & Gold, 1972). However, black and white adolescents differ in the types of crime they report. Blacks are much more likely to report violent or more serious crimes whereas whites are more likely to report certain property or less serious crimes. To some extent, blacks may be disproportionately represented in official reports because the type of crime they commit is disproportionately represented. However, since the evidence is mixed, ethnicity may still be considered as a source of bias separate from the seriousness of the offense.

Gender is another source of potential bias in official records. While self-report studies agree with official records that women are less frequent law violators than men, the two measures disagree to some extent, about what type of law violation each gender commits. Some studies (e.g., Hindelang, 1971; Williams & Gold, 1972) suggest that although girls commit the same range of crimes that boys do, the crimes are not committed with the same frequency by both genders. For example, boys report committing more violent crimes than do girls. For some crimes, however, the self-report rates are equivalent while the arrest rates differ. Teilmann & Landry (1979) report that girls are overarrested for status offenses such as runaway and incorrigibility. Thus girls would be more likely to appear in arrest records for less serious crimes, although differences in behavior may be small.

A final consideration about the person that may be a source of bias in official records is the frequency with which the individual commits crime. William and Gold (1972) found a modest positive correlation between the probability of arrest and frequency of offense. Furthermore, it is well known that prior arrests are the best predictor of future arrests. Thus, the one-time or infrequent offender is likely to be underrepresented in official records.

Self-report of criminal activity is not without its biases and short-comings. Error can be introduced both as a function of the task itself and as a function of the technique used to elicit self-report. The self-report method asks the individual both to recall and to reveal criminal activity. Individuals will vary in their ability to recall beyond the recent past. Furthermore, recall may depend on the salience of the criminal act which may be affected by the seriousness of the criminal activity and by the individual's own comparison level for criminal behavior. For example, a serious distant past crime such as burglary may be more salient to an individual than a recent incident of driving under the influence of alcohol. And, if the individual chronically drives under the influence, he or she may be less likely to recall a particular incident than a less frequent occurrence of petty theft. Thus, bias in self-report may enter at the level of recall.

Bias may also come from differences in motivation to reveal crime. Some studies (e.g., Gold, 1970; Farrington, 1973) have shown that serious crime is underreported in self-report studies. Assuming that serious crime is more salient and memorable, we may conclude that people do not want to report accurately serious crimes they have committed.

To some extent, accurate reporting will depend on the technique used to obtain self-report. Teilmann (1977) has shown that a match in race and gender between interviewer and reporter tends to enhance the amount of reporting.

Interviews, however, are time-consuming and costly. An alternative technique, questionnaires, suffers from low return rates and the limited ability of some respondents to respond adequately to written questions. All self-report techniques allow refusal to participate which may introduce

its own bias if refusals are not random. Thus, for many research questions, in consideration of the biases in self-report, its higher cost in time and its lower accessibility may make it less appealing than official records.

Since many of the biases of self-report involve the type of offense reported, it is important to know how well official records estimate the true population distribution of offenses. Hindelang (1974), in a comparison of the Uniform Crime Reports to a large victimization survey, found that, after adjusting for differences in absolute volume of offenses, the order of seven offense categories was the same across the two studies. Furthermore, the relative frequencies in five of the seven categories were virtually the same. Only burglary (underreported in UCR) and vehicle theft (overreported in UCR) showed major divergence.

If a researcher chooses to use official records, he or she must decide whether to pursue those records at a local (city or county) or statewide level. One advantage of using local records is that particular types of offenses may not be reported to a statewide agency. Some state-wide agencies require a certain level of seriousness before they will accept the report of a crime. In addition, local agencies, if given the option, may not take the trouble to report less serious offenses. Thus, the chances of finding a less serious crime may be better at the local level.

Another advantage of local records is that generally charge and disposition information for a particular arrest is more likely to be complete than at a statewide agency.

A disadvantage of local records is that they mask the mobility of an offender. A search of statewide records would tell the researcher whether an individual had been arrested anywhere in the state rather than in only one local area of the individual's home address. Moreover, if specific home addresses are not available, statewide records may be the only place that a search can be made. If the quality of charge information, the disposition information or the seriousness of the charge will affect the research issue, the statewide system should be examined for its completeness on these variables. Although there may, in general, be a tradeoff between quality of information and breadth of coverage, some specific statewide systems maintain high quality information.

The remainder of this chapter will describe statewide record systems in a number of states, outlining particular features and general caveats in the use of these statewide records.

Available Resources

Seventeen states were surveyed about a statewide central arrest record system and the possibility of research access. The seventeen states represent the geographical location of most of the longitudinal data bases described in this book. In all states, a responsible party in the appropriate state agency was interviewed. Additionally, academic researchers in five states were contacted for their opinions about criminal research in their state. Their opinions will be integrated into different discussions of what is available in each state.

A list of the seventeen states' agencies and responsible individuals for criminal history records was supplied by the Bureau of Criminal Statistics (BCS), California's central repository for arrest reports. This list was supplemented by information from SEARCH group, a Law Enforcement Assistance Administration-funded organization that compiles information on computerized criminal information systems. SEARCH group, based in Sacramento, California, publishes a biennial Directory of Automated Criminal Justice Information Systems which can be obtained from LEAA in Washington, D. C. Interim updates may be obtained from SEARCH group in Sacramento.

Neither BCS nor SEARCH group knew of systems in Vermont or Washington state. A Vermont state sourcebook suggested a central record system for criminal histories, which was subsequently contacted. According to the BCS contract, the State of Washington was to begin a central repository for Uniform Crimes Statistics in 1979. Prior to that the individual police agencies in Washington reported directly to the FBI. The BCS contact, who was consulting on the development of Washington's centralized system, speculated that centralized criminal history information will not be available at the state level until at least 1985.

Those surveyed about states were asked questions covering the following areas:

I. Composition of the system

   a. What types of offenders?

   b. Does the system keep records on juveniles?

II. Information available about individuals in the system

   a. What information can be found on a typical file?

   b. How long have records been kept?

III. Reporting procedures and record keeping

    a. Who reports to the system?

    b. How frequently are reports made?

    c. Is the system computerized or manual?

    d. What changes, if any, are planned for the system?

IV. Access

    a. What are the possibilities and avenues for access by researchers?

    b. What information would a researcher need about an individual to find him/her in the system?

Table 1 summarizes answers to the questions above from each of the 16 states with record systems where there was variability across states. Those question areas that were answered much the same by all states will be presented in the general discussion of the question areas that follows rather than in Table 1.

A word of caution must be issued about the information in Table 1. The information presented is the authors' interpretation of answers given by usually only one representative of the state agency. Thus, the information presented may not completely accurately portray what is available in a system for a particular research study. Researchers interested in accessing a particular state system are encouraged to communicate with that system about the areas covered by the survey with respect to a specific research plan or question.

Composition of the System

Most states keep criminal history files on individuals who have been arrested and fingerprinted. Vermont and California are exceptions. In Vermont, anyone who has been arraigned for a criminal offense is reported to the state system. California receives a report on anyone arrested. Each state's reporting agency determines what it considers an arrest. In some agencies, booking is necessary for an arrest whereas in other agencies, a field citation will be counted as an arrest.

The states that record fingerprintable offenses vary in what offenses they fingerprint. In most states, felonies and major misdemeanors are fingerprintable. Some states, such as New York, Florida, Pennsylvania, Illinois and Michigan fingerprint for minor misdemeanors. West Virginia and Michigan base the decision to fingerprint on a specified level of punishment for the offense. Fingerprints in Michigan are mandatory on all offenses except misdemeanors for which the punishment is less than 90 days in jail or $100 fine. Since criminal codes are defined at the state level, states may also vary in what offenses are considered felonies,

major misdemeanors and minor misdemeanors. Although listing fingerprintable offenses by state is beyond the scope of this paper, it would be beneficial for the researcher to identify the pool of specific offenses in the state system of interest.

In some states, that an offense is fingerprintable does not necessarily ensure that all offenders will be fingerprinted. States such as Texas, North Carolina, Iowa and Pennsylvania (as of 1979) have voluntary fingerprinting. Respondents in these states however, believed that fingerprinting of felonies was very frequent (around 90%).

That states vary in who is reported to a central record system may not matter for many research questions with longitudinal data bases. Of more importance is whether a state keeps its rules constant over time for entry into the criminal record system. Unfortunately, perfect constancy cannot be reasonably expected because states may change their laws about what is fingerprintable, whether certain crimes are felonies or misdemeanors and whether fingerprints are mandatory or voluntary. For example, in New York prior to 1978, possession of marijuana was a felony and, therefore, fingerprintable. After 1978, it became a minor misdemeanor and disappeared as a sole cause for entry on the state system. Pennsylvania, in 1979, had voluntary printing but passed a law to make printing mandatory for felonies beginning in 1980. Such a change may have little impact, however, if the voluntary printing of felonies is as high as claimed.

The composition of the records kept in each state may influence which state is chosen to explore certain research questions. Research studies that wish to maximize the probability of finding any criminal activity will be better served by states such as Illinois and Michigan which include a wider range of offenses than states such as Texas and Florida. Alternatively, studies that need only to find felony bookings would be about equally successful in all the surveyed states.

The record systems surveyed generally included only adults, i. e., those 18 years and older. If juveniles are recorded in a state's system, access to their files is reported to be almost impossible for anyone. Access in California is better than in most states. Wisconsin and Nebraska keep indexes of runaways at the state level but, again such indexes are said to be difficult to access. Four states (Texas, Florida, New York and California) are developing or improving a computerized record system on juveniles incarcerated in the state juvenile institution. The system used by these states is called Child and Youth Center Information Systems and provides a summary of the juvenile's commitment history, various adjudications, demographics, family background and drug use. As Empey (1978) has pointed out, the group of juveniles found in a system such as CYCIS tends to be serious, chronic delinquent offenders. These juveniles are a small subset of all juvenile offenders. However, even though CYCIS represents a select group of juvenile offenders, it will be accessible for researchers and may prove useful for some research interests.

Information Available

Two areas were discussed in the survey: 1) what comprises a typical criminal history file and 2) how long have records been kept.

All states surveyed record date of arrest, charges, identifying characteristics such as sex, date of birth and address, state identification numbers, and disposition for each arrest submitted to the system. Every state except California said that they included all arrests received on an individual in that individual's file. California keeps records by arrest rather than by individual, consequently a complete arrest history would have to be searched and compiled by the researcher for each research subject. Specifically, in California, records from each year would have to be searched to see if the individual in question had been arrested.

Disposition information in most states was reported as the least likely piece of information to be available about a particular arrest. Two states, North Carolina and Wisconsin have mandated audit systems that check all files for information completeness primarily for courts and disposition. Where information is missing, it is obtained from the relevant court. Pennsylvania expects to institute such an audit system by 1981.

States that have Offender Based Transaction Statistics (OBTS) systems, such as Illinois, California and New York, hold promise of improved completeness of disposition information when the systems are fully operational. OBTS systems are being instituted in various states through funding from LEAA. These systems allow the individual to be tracked from arrest through to each case's final disposition.

Completeness of disposition information may be of little concern to some researchers who only need arrest information. However, as of 1979, Nebraska would not release any file on an individual unless disposition information was complete. No other state surveyed mentioned such a restriction, although, in keeping with privacy concerns, other states may in the future decide to impose such a restriction.

Starting dates for the states' record systems are presented in Table 1. Where two dates are displayed, the later date usually represents the beginning of a computerized system. Computerized systems will only be accurate in prior histories from the year in which the system was begun unless there has been some activity on an individual's file. For example, if the system began in 1972 and an individual was arrested during or after 1972, his or her entire computerized prior record will include everything found in paper records (many times going back to the 1930s). If, however, the individual was arrested before 1972, with no subsequent record in the system, the individual is likely either not to appear in the computerized system or to appear on a computerized name index only, in states where such an index is kept.

Reporting Procedures and Record Keeping

Law enforcement agencies report directly to the agency listed in Table 1 in all 16 states except Massachusetts. In Massachusetts, as of 1979, arresting agencies report to the courts. Probation Departments keep these records. Thus, the central repository in Massachusetts is the Office of the Commissioner of Probation. Access is obtained through the Criminal History Board.

The frequency with which law enforcement agencies submit arrest information to the central system affects the likelihood of finding recent arrests in the system. If a researcher is interested in as many arrests as possible, systems that require reports within 24 hours, such as New York, Illinois, New Jersey and Wisconsin, will provide optimal opportunity. Other systems that have delays in reporting (e. g., California) may contain only arrests a month old or more. The interval between arrest and receipt of information about the arrest by the central agency is reported in Table 1 for each state surveyed.

Whether or not the record system is computerized is also reported for each state in Table 1. A state is considered computerized if both name and arrest information can be obtained from the computer. A state is considered not computerized if paper files must be manually searched for all information. One state, Wisconsin, has only a name index computerized. Computerized files make data collection faster and less expensive. However, as an informant in one state embarking on a computerized system commented, information loss may be increased with computerized files.

Access

Perhaps as important as the quality of information in an official system is whether a researcher can gain access to that system. Even the most comprehensive system is useless to the researcher if it is not accessible. All criminal record systems are governed by federal privacy regulations that restrict the dissemination of criminal information with identifying characteristics. States may also have laws regulating access to criminal records. Federal law and most state laws allow the use of criminal records for statistical purposes. A researcher using criminal records must agree to accept the data free of identifiers and to guarantee that any published results will not allow for any individual with a criminal record to be identifiable.

Researchers using a pre-existing longitudinal data base would probably want to search the record systems with names and other identifying characteristics of those in the data base. Since such a search emphasizes identifying characteristics, it would be carefully scrutinized for compliance with privacy regulations. Most states surveyed said they would consider requests to make such a search on an individual basis, evaluating the merit of the research issue to be studied. Requests should be addressed to the agencies listed in Table 1 for the relevant state, except where a separate agency is listed in the Access column.

Of those surveyed, four states placed severe restrictions on access by researchers. West Virginia and Vermont replied that their systems would be very difficult to access. Vermont allows access only through law enforcement agency teletypes. Pennsylvania said that state law, as of 1979, prohibits access to anyone except police, district attorneys and Commonwealth administration. However, more open access in Pennsylvania was anticipated by the enactment of a new law in 1980. Nebraska will release to law enforcement agencies any record. However, it will release to the public (which includes researchers) only cases that have recorded final dispositions, i. e., cases that have not been completed or for which there is no record of completion are not made available to researchers.

Those cases without recorded dispositions represent a nontrivial proportion of all cases in the system. Therefore, Nebraska, in effect, allows only limited access. No firm plans are in motion in Nebraska to increase researcher access, although the issue is being considered by the State Judiciary Committee.

In contrast to Nebraska, Florida allows public access to any criminal record that has not been sealed. However, it was reported that Florida is working on legislation to insure more privacy of the records.

Two states reported that a fee would be charged for each name searched in their record system. Florida imposes a fee of $2.00 per name which may be waived for researchers. New York charges everyone allowed access $5.00 per name when the state identification number is not attached at arrest and $2.00 per name when the state identification number is attached.

The basic information needed to locate an individual in any of the state systems is name and date of birth. Some states, such as Florida and New York, operate soundex name indexes which will produce for each name requested a list of those names with different spellings that sound like the requested name. Such a system allows for leeway in the recorded spelling of a name and increases the chances of finding the individual in the system. However, since many names may be produced for each name searched, New York imposes the restriction that no information can be released on a case unless the correct case is identified with certainty. Thus, in New York, additional identifying information such as address, ethnicity, gender, height or Social Security number may be necessary to access a criminal history file.

Discussion

Once the researcher has located an appropriate statewide system with adequate information and reasonable access, he or she should continue to be aware of the possible problems in using the system. This paper will conclude with a general discussion of two types of problems one may face in using official records.

The first problem concerns the researcher's inflated expectations after first finding a viable records system. When discussing what is available in a state record system with someone in that system, the researcher may be overly optimistic about the project because everything the researcher needs to know is allegedly available in the system. However, once the researcher begins to collect data, he or she will likely discover that what was thought to be available is not easily obtainable. For example, the California BCS system will provide to an approved project data with the second through sixth characters of the individual's name which can then, with reasonable statistical certainty, be used to identify the particular individual. However, the location of these "identifying characters" on the person's record varies from file to file. Thus, the process of obtaining identifiers becomes very complicated.

Similarly BCS officials assert that they can provide criminal history information for an individual. The information is available but must be compiled by searching each year's files for the particular individual. If the researcher had assumed that, as in most states with OBTS systems, California OBTS compiled criminal history information, he or she would at some point be faced with unexpected time and cost requirements.

In one final example of mismatched expectations, one group of researchers came to believe, based on information from a central records bureau, that they had the definitive list of a set of juveniles committing certain serious crimes for a particular year. However, when the researchers went to check paper files on the listed juveniles, they discovered many had not been charged with the crimes the central agency recorded for them. This calls attention to the problem of overconfidence in the information provided by the records bureau.

A second type of problem with statewide record systems and any official record system is the system's tendency to change over time. Many times systems will change in an effort to improve the quality of record-keeping and retrieval of information. Changes from manual to computerized files (e. g., Iowa) or from one computerized system to another (e. g., New York) may require that the format in which the information is reported also change. Such format changes, even when mandated by law, rarely are put into effect simultaneously by all reporting agencies. Thus, in comparing data from several years, large chunks of information may be missing due to report format changes.

An example of such a format change took place in California's BCS. In 1974, 61 percent of the arrests reported to BCS were in the proper format for automated record-keeping. By 1977, 95.4 percent of the arrests were reported using the automated format. For the data sets to be comparable across the four years, jurisdictions entering the system late would have to be cause for elimination for data from all years. In the example given here, this would represent a major loss of information resulting from an improvement in record-keeping.

Another way that a statewide record system may change over time is through the inclusion or deletion of agencies that report to it. In California, for example, the BCS/OBTS system was started in southern counties in 1975 and did not receive information from all counties until the end of 1977.

Similarly, two large southern California Sheriff's departments were not sending arrest information to BCS until late 1976 and early 1977. Thus, at one point in time the data base for criminal records is quite different than at another point in time. This difference may confound a study's results, particularly if the study concerns an analysis over time.

Quality of reporting can also be affected by changes in political leadership. Booth et al (1977) suggest that more complete reporting from local to state level occurs when liberal politicians are in power. If such a difference in completeness of reporting exists, it may cause problems for certain analyses, such as in comparing two cohorts formed under different political leaderships.

In addition to internal changes in the record system to improve quality, outside changes, usually through laws, can affect the composition of the records system. One such legal change which affects what arrests are recorded (changes with fingerprinting requirements in Pennsylvania) has been discussed earlier in this paper. However, another type of legal change may also contribute dramatically to a change in the composition of the records systems. This type of change affects the arrest rates for a particular offense, usually through an alteration of the penalty. For example, the penalty for possession of marijuana in many states has been reduced from incarceration to a fine. Even though possession of marijuana is still a crime, arrest rates drop with such a reduction in penalty. Comparing records from before and after such a change in law may lead to spurious conclusions stemming from the law rather than the variables of interest.

Although changes in the record systems may create problems for insuring adequate controls in research designs, change, particularly that aimed at improving record systems, may also be beneficial to the researcher. In the future, we may find systems that are more standardized, accurate and comprehensive than those available today. With improved record-keeping and continued access, researchers may, in the future, find that official records in the United States are the most viable measure of criminal activity.

Ease of Access to Criminal Registers by States
(for states with longitudinal projects)

State	Agency-Address	Who is Registered?	Juveniles Registered	Computerized?	Access	When Started	Speed of Report
California	Bureau of Criminal Statistics, Sacramento	All arrests	Yes	Yes	By application	1975	Monthly
Florida	Florida Crime Information Bureau, POB 1489 Tallahassee	18+, fingerprints	Some, no history	Yes	Yes, with fee	1969	1-7 days
Illinois	Bureau of Identification, Department of Law Enforcement, Division of Support Systems, 300 N. Armory Bldg. Springfield, 62706	17+, fingerprints, felonies most misdemeanors	No	Yes	Offender Based Transaction Statistics information released only. Individual requests considered.	Some 1932, All 1976	Within 24 hours
Iowa	Offender Based State Correctional Information Division of Adult Corrections, Department of Public Safety, Hoover State Office Bldg., Des Moines	Fingerprints and point system based on charge	No	No	Individual requests considered		Weekly, some yearly
Maryland	Criminal Records Repository Maryland State Police Headquarters Pikesville, 21208	Any criminal arrest (Maryland Article 27)	State Police arrests only	Yes	By petition to Secretary of Public Safety	Some 1940, All 1978	Within 72 hours

Ease of Access to Criminal Registers by States (continued)
(for states with longitudinal projects)

State	Agency-Address	Who is Registered?	Juveniles Registered?	Computerized?	Access	When Started	Speed of Report
Massachusetts	Criminal History Board, 1 Ashburton Place, (21st floor) Boston	18+ finger-prints	none	No	By appli-cation to Criminal History Board	1920	After arraign-ment
Michigan	Criminal Justice Data Center, Michigan State Police, 714 S. Harrison, E. Lansing, 48823	Fingerprints, felony, and most misdemeanors	Some 17yr-olds, otherwise, no.	Yes	Individual requests to Commander, State Police Central Record Bureau	Some 1930, All, 1975	1-2 weeks
Nebraska	Commission on Law Enforcement and Criminal Justice POB 94946, Lincoln 68509	Fingerprints, felonies, some misdemeanors. file purged if no activity in 15 yrs.	Escapees from state facilities, runaways	Yes	Limited now, even with access. A record cannot be released unless complete. (i.e. include final disposition)	Some 1943 All 1972	Within 15 days
New Jersey	Records and Identi-fication Bureau Division of State Police, PO 7068 West Trenton, 08625	Fingerprints, any indictable offense (excludes most misdemeanors)	No	Yes	Contact Divi-sion of Systems and Communications, need authorization of State Attorney General	1972	Within 24 hrs.
New York	Division of Criminal Justice Services, Executive Park Tower, Stuyvesant Plaza Albany, 12203	Fingerprints, felonies, most misdemeanors	Yes	Yes	By application to Division of Criminal Justice Services, with fee for search	1973, prior dates available but not reliable	Immediately upon arrest by Laser-Fax

Ease of Access to Criminal Registers by States (continued)
(for states with longitudinal projects)

State	Agency-Address	Who is Registered?	Juveniles Registered?	Computerized?	Access	When Started	Speed of Report
North Carolina	Police Information Network, PO 27047 Department of Justice, Raleigh, 27611	16+, finger-prints, felony and major mis-demeanors	No	Yes	Address request to State Bureau of Investi-gation	Some 1950, All, 1974	1-2+ days
Pennsylvania	Bureau of Research and Development Records and Identi-fication, Pennsylvania State Police, 1800 Elmerton Avenue, Harrisburg	60-70% criminal arrests; finger-prints, 18+ felony	Some from large cities if finger-printed	No	None, how-ever, law going into effect in Jan, 1980 may open to researchers	1924	1-5 days
Texas	Texas Crime Infor-mation Center Department of Public Safety, POB 4143 Austin, 78765	Fingerprint, 90% felonies 80% major misdemeanors 17+ (voluntary reporting)	No	Yes	Consider on indivi-dual basis, send requests on to Director	1970 if before name index	1-7 days
Vermont	Vermont Crime Information Center Montpelier	Any criminal offense that has been arraigned	Yes	No	Through Law Enforcement Agencies only	1947 at State Police 1970 at Vermont Crime Information Center	After arraign-ment
West Virginia	Criminal Indenti-fication Bureau, Department of Public Safety 725 Jefferson Rd. South Charleston	Anyone committing crime for when confine-ment is punishment, fingerprints	Yes, if finger-printed	No	Very diffi-cult	1935	1-7 days

Ease of Access to Criminal Registers by States
(for states with longitudinal projects)

State	Agency-Address	Who is Registered?	Juveniles Registered?	Computerized?	Access	When Started	Speed of Report
Wisconsin	Crime Information Bureau, Department of Justice, POB 2718, Madison	18+, finger-prints (felony and some misde-meanors)	Yes but separate files	Part, Yes for Name Index, no for transcripts	By appli-cation to Bureau Director	1971	Within 24 hours

References

Booth, A., Johnson, D. R. & Choldin, H. M. Correlates of city crime rates: Victimization surveys vs. official statistics. Social Problems, 1977, 187-197.

Empey, L. T. America delinquency: Its meaning and construction. Homewood, IL: The Dorsey Press, 1978.

Farrington, D. Self reports of deviant behavior: Predictive and stable? Journal of Criminal Law, Criminology and Police Science, 64; 99-110, 1973.

Federal Bureau of Investigation. Crime in the United States: Uniform Crime Reports, 1975. Washington DC. U S Government Printing Office, 1976.

Gold, M. Delinquent behavior in an American city. Belmont CA: Brooks/Cole Publishing, 1970.

Hindelang, M. J. The Uniform Crime Reports revisted. Journal of Criminal Justice, 1974, 2, 1-17.

Hindelang, M. J. Age, sex and the versatility of delinquent involvements. Social Problems, Spring, 1971, 18: 527-535.

Hirschi, T. Causes of delinquency. Berkeley: University of California Press, 1969.

Skogan, W. Dimensions of the dark figure of unreported crime. Crime and Delinquency, 1977, 23, 41-50

Teilmann, K. Sources of bias in self reported delinquency. Unpublished dissertation. University of Southern California, 1976.

Teilmann, K. & Landry, P. Gender bias in juvenile justice. Journal of Research in Crime and Delinquency, in press (1981).

Williams, J. R. & Gold, M. From delinquent behavior to official delinquency. Social Problems, 1972, 20, 209-229.

Wolfgang, M. E., Figlio, R. & Sellin, T. Delinquency in a birth cohort. Chicago: University of Chicago Press, 1972.

Chapter 10

Confidentiality and Ethics

in

Longitudinal Research

Michele Harway, Ph.D.

University of Southern California

Social Science Research Institute

CONFIDENTIALITY AND ETHICS IN LONGITUDINAL RESEARCH

Ethical issues in the context of social science research present a problem stemming from the tension that exists between two sets of related societal values. These are: a belief in the value of free scientific inquiry, and, at the same time, a belief in the dignity of humankind and the right to privacy. Reconciling these two frequently contradictory values is a problem that a researcher must face each time a new study is undertaken because it is impossible to delineate a specific set of rules and regulations governing all research studies. In each instance, the researcher must decide on a course of action after giving careful consideration to the importance of the study and the extent of potential injury to the dignity of the participants.

In this chapter, we will review the major ethical constraints faced by longitudinal researchers. This is not to say that many of the ethical issues reviewed here will not also impinge on researchers in other areas of the social and behavioral sciences, just that our discussion will be specifically oriented to longitudinal research. It is likely that many more questions will be raised than queries answered, for ethics, except where controlled by laws, are a matter for the individual conscience.

Individual Rights to Privacy

One of the obvious implications of longitudinal research is that respondents or subjects in such studies are subjected to scrutiny over extended periods of time, thus potentially interfering with these individuals' right to privacy for months or years. However, the issue of privacy evaporates if subjects are fully informed and freely consent without coercion to participate in a research project, as is the case in most longitudinal studies. Thus, informed consent is usually considered as one element in the protection of subjects. In many cases, however, truly informed consent can be difficult. One reason for this is that it is not always possible for the researcher to explain the research procedures to participants, particularly in studies which involve complex variables that are difficult to explain succinctly. Moreover, naturalistic observation of groups must sometimes be made unbeknownst to subjects. Obtaining informed consent is also not feasible in other situations: In Great Britain and the Scandinavian countries where the longitudinal tradition is well established, many longitudinal projects have been made possible because of the rich source of data provided by the national registers. Until recently, it has been possible under certain circumstances to utilize these data for research purposes without informed consent of the subjects. This could be considered by some an ethical oversight; yet it is clear that incorporating these data into longitudinal projects has yielded invaluable results for mental health, social policy, and medicine. In this country, of course, obtaining equivalent U.S. data (e.g., census data) is not possible because of a special public law which forbids access to this information. If American researchers avail themselves of the Scandinavian data files for research purposes, they must still weigh for themselves the ethical issues involved.

Another area where the issue of informed consent presents a serious problem for longitudinal researchers is in considering the ethical responsibility to subjects whose informed consent extended only to the original data

collection, not to subsequent follow-ups. This issue will be discussed in more detail in a later section.

Confidentiality

Another element in subject protection is confidentiality. The researcher has the responsibility either to collect data anonymously or to provide confidentiality both in public reports and in research records. Anonymity and confidentiality are related. Anonymity refers to preventing the identity of the respondent from ever being known. Anonymity merges with confidentiality when the research design permits the identity of respondents to be known at a point in time to a limited number of investigators but is otherwise protected from dissemination.

The question of confidentiality is of particular concern in longitudinal research since collecting anonymous data is not usually possible: With this technique it is critical to maintain identifying information on respondents so that they may be followed up during the multiple data collection waves that comprise most longitudinal projects. Usually, not only is name and address information kept for all respondents, but since follow-up itself is often a problem (see Ch.), other relevant information that would allow for the tracking of individuals is maintained. Thus, some project staffs collect names and addresses of relatives and close friends, respondents' driver's license numbers or social security numbers, all in attempts to provide leads to the whereabouts of a cohort member who has moved from the last known place of residence. Confidentiality presents a greater problem to the extent that more identifying information is kept on respondents. Longitudinal researchers have to be particularly sensitive to this ethical issue.

It may be useful in discussing the concept of confidentiality to consider the points at which confidentiality could potentially be violated. Discussed here are four points in the data collection and analysis process where such violations could occur. The four points are: 1) Data collection 2) Data processing 3) Physical document storage, and 4) Statistical data storage. There may, of course, be other points where violations are also possible.

Prior to data collection is, of course, the time when initial decisions are made about whether to collect data anonymously. In the case of longitudinal research this is not usually a viable option. But consider the following: Not all longitudinal research involves analyses which compare individuals to some earlier baseline data. Rather, in most cases what is of interest is to compare subgroups of respondents with certain unique characteristics to other subgroups of respondents. In those instances, it would not be necessary to include the name and address of the respondent directly on the questionnaires or other data collection instruments. In fact, the only reason for keeping names and addresses in such instances is to ensure that the same individuals are followed up consistently in later data collection waves. Follow-up can be easily accomplished if a separate name and address file is maintained from the beginning of data collection. Then a procedure such as the postcard technique (used successfully by this author) can be called on to keep track of who is responding to the quest for information and who needs to be contacted further.

The postcard technique is fairly simple. It consists of sending to each respondent (in the case of a mail questionnaire) an anonymous questionnaire

and a postcard (return-addressed and stamped). The reverse side of the postcard lists the subject's name and address and requests that he or she check one of the two boxes ("I am returning the questionnaire" or "I do not want to participate further") and return the card separately but at the same time as the questionnaire. In such a way, the investigator can keep track of who is returning the instrument and who is refusing participation. The researcher also knows whom to send a follow-up to: subjects from the original mail list who have not returned a postcard. At the same time, questionnaires returned are completely anonymous. Needless to say, this technique will not work in studies where some link-up of the data will be required.

A second step in the research process where confidentiality may come into play is at the processing stage where data is coded or organized. Assuming the typical longitudinal study where identifying information on the respondent is maintained as part of the file, then at this stage the respondent's confidentiality could be violated. Of course, having trustworthy staff members is a critical and obvious element. Moreover, it is possible at this stage to implement a procedure to separate identifying information from the original data forms while retaining the data elements and allowing for their processing. Thus, it would be possible to set up a name and address file which would also include a respondent identification code. Then, actual respondent identifying information could be destroyed from the original data forms (shredded in a paper shredder if found on a separate page from the data or inked over if on the same page).

The third point at which confidentiality can be protected is in the storage of the original data forms once processing is complete. In cases where it is critical to keep the original data collection instruments (as in interviews which were coded for specific hypothesis testing but which might yield further information if recoded to test a different set of hypotheses), it is critical that these documents be kept in a secure location. If separation of identifying information has not yet occurred, then it should be considered at this stage. Where names and addresses have been culled from the original data forms and set up in a separate file system, then it would be desirable that these files be physically separate and secure to be accessible for purposes of follow-up.

In cases where processing of the physical documents has involved computer scoring of answer sheets or keypunching of precoded instruments, it may not be necessary to retain the physical documents themselves. It would be sufficient in those cases to maintain a separate name and address file and a set of the keypunched or computer-scored data. If the physical documents are going to be disposed of, then it is important that their disposal be accomplished within the bounds of the confidentiality concern. That is, if names and addresses remain on the questionnaires, putting them out with the trash is a dangerous procedure. Shredding or burning should be considered.

The fourth step in the research process is that of statistical data storage. As in the storage of the physical data, separate storage of the statistical data and identifying information is also recommended. An additional precaution would be the development of a controlled data access system making it impossible for any users of the longitudinal data, including staff members, to obtain any information about the identity of individuals, but continuing to allow full access to the data itself. Such a system was developed by the American Council on Education (ACE) in the late 60's to protect the rights of their research subjects.

In order to understand the development of the ACE's Link file (as it came to be known), it may be necessary to know something of the political climate which reigned at the time. First, however, it is important to state that concerns about confidentiality and the ethics of research are relatively new. The ethical tenet of protecting the confidentiality of data has by now been incorporated in courses on research methodology and figures prominently in the ethical codes of organizations such as the American Psychological Association and other professional groups. Human Subjects Protection Committees on university campuses, in the administration of research institutes and in government funding offices scrutinize research grant and study applications to ensure that confidentiality is adhered to in the collection and storage of research data. Yet, this heightened awareness of the importance of confidentiality was not always in existence. As recently as 1967, the Executive Office of the President judged these issues to be of sufficient importance that a panel of distinguished scholars (headed by Kenneth E. Clark) was convened to prepare a report on privacy and confidentiality in behavioral research. Up until the publication of that report, these ethical constraints had not received much attention. Shortly thereafter, concern about issues of confidentiality manifested itself within a number of professional associations in the social sciences. Thus, the 1968 American Personnel and Guidance Association Convention featured a symposium which addressed issues of confidentiality of data, respondent rights and administrative problems in making data available to researchers. The American Educational Research Association formed a committee in 1969 to evaluate alternative approaches to existing problems arising from concerns about confidentiality. The Russell Sage Foundation demonstrated its concern by funding several projects on privacy and confidentiality. As stated earlier, some professional associations, notably the APA, incorporated guidelines for dealing with confidentiality into their ethical principles.

In most cases, while ethical codes were put forth, no procedures were suggested for implementing these. One notable exception was the American Council on Education's (ACE) link file. Yet, even that file was developed under duress. In 1969, ACE was charged by NIMH with doing a study of campus unrest at the height of the student protest era. Shortly after the beginning of that study, protest about the study itself sprang up (the interested reader is referred to Science, 11 July, 1969, for a description of the events surrounding that protest). The crux of the concerns revolved around fear that the data collected for scientific purposes would be used by governmental authorities "for punitive action against individuals" (Science, 11 July, 1969 p. 157). This concern would involve a clear violation of the privacy and confidentiality of the respondents. Guidelines designed to insure the confidentiality of study data were also criticised on the grounds that: "If any government agency subpoenaed data, the ACE research office would be truly in the middle and would face a test of a reported assurance to critics by ACE staff that they are 'prepared to go to jail if necessary' in order to keep the data confidential" (pp. 158-159). In the paranoia of the times, there was concern, perhaps legitimate, that social science researchers would not take seriously their ethical responsibilities when a government subpoena for the identity of their subjects was involved. ACE's eventual solution to the dilemma was the development of a link file.

The ACE link file is conceptually quite simple. It consists of three files for every set of data. The first file is a name and address file for

all members of the cohort. The names and addresses are paired in each case with a set of identification numbers. The second file consists of the data for the cohort paired up for each subject with a second set of identification numbers unrelated to those maintained in the name and address file. The final or link file consists only of the two sets of numbers: The identification numbers from the name and address file for each respondent paired up with the identification numbers for each respondent from the data file. The data file and the name and address file remain on the premises, permitting both easy access to the data and access to a complete list of respondents for follow-up purposes. At the same time, the link file is deposited in a computer facility located in a foreign country, making that file inaccessible to American subpoenas. The final safeguard in such a system is an agreement with the foreign computer facility that "they will neither copy the file nor make it available to outside persons, including research personnel of the American Council on Education. The foreign facility is bound to this agreement even in the event that the American Council on Education should subsequently request that the file be returned... the foreign facility is under no circumstances to release this link file to other individuals or organizations. Thus, both ACE and the foreign agency must violate the agreement before research data can ever again be matched directly with identifying data." (Astin & Boruch, 1970, p. 61)

Linking up follow-up data with the original data collected on the cohort can be a bit tedious when this system is used, but the additional protection that it offers may make it worthwhile. Linking up data involves a 6-step process. The first step in a follow-up involves printing address labels from the name and address file. These labels also include the identification numbers located on that file. Step 2: When the completed questionnaire is returned, it is processed and the responses generated by the instrument and the identification number are transferred to computer tapes (no names and addresses are transferred). The original questionnaires are then destroyed in Step 3. In Step 4, the tape is sent to the foreign computer facility where it is copied with the ID number replaced by the second one. The new file is then sorted on the second number in Step 5 in order to put the records in a different order. In the sixth and final step, the newly processed file is then returned to the researchers for merging with the original data set.

Informed Consent and Secondary Data Analyses

With this volume we are trying to encourage individuals who are interested in longitudinal research to consider becoming involved in collaborative utilization of existing longitudinal data files. As we discussed in another chapter, not only is longitudinal research time-consuming, since it requires many years to build up a longitudinal file, but also funding agencies are becoming increasingly reluctant to fund large multi-year longitudinal data collections. Thus, secondary analyses of existing longitudinal data files to test research hypotheses not of interest to the original investigator make ultimate good sense. However, conducting such secondary analyses brings up an ethical question related both to confidentiality and informed consent.

The ethical question is the following: When an individual consents to participate in a longitudinal project after being informed of the study's procedures and goals and after being guaranteed confidentiality, what are the rights and responsibilities of the original investigator and of a second investigator who is doing secondary analyses with the data? Our discussion of this

issue will focus on two cases of informed consent: one where the respondent has agreed through a signed consent form to participate in the research project, and the other where there is implicit agreement to participate but no contractual document.

The first instance, that of a signed consent form, is still a relatively rare occurrence. Federal agencies such as the Department of Health, Education and Welfare now have regulations regarding the protection of human subjects, including the requirement of written consent forms. However, this is a relatively recent phenomenon and longitudinal projects which were initiated prior to the 1970's almost certainly did not require that respondents sign such an agreement. For new projects, however, where a consent form has been procured, does the original agreement between the respondent and the first researcher cover subsequent investigators as well? This is a difficult question to answer, especially in situations where the subsequent researchers are not and have never been on the staff of the original investigator. Technically, the contractual agreement was between the original investigator (and his or her staff) and the respondent. Consequently, should another investigator at a subsequent time wish to conduct follow-up research on this same cohort without working collaboratively with the original investigator, then permission should probably be secured from the respondents. In most instances, however, this kind of situation is unlikely, since the original investigator would probably remain involved with the project in some capacity. When the involvement of a second researcher consists only in conducting secondary analyses on a data file where individual identities are not available, then no ethical violations would be incurred. No ethical problem would be involved either in a situation where anonymous data were being analyzed. In both instances, since the individual's right to privacy and consequently his or her confidentiality was not endangered, this would not present an ethical issue.

Where informed consent was an implicit agreement between respondent and researcher, the same guidelines would apply. This is particularly the case when a subsequent investigator would like to complete a follow-up of the cohort independently of the original researcher. In those situations where the original researcher was no longer involved in any capacity in the project, then it would seem incumbent on this latter researcher to contact the individual cohort members and secure written consent to participate in the follow-up, even though no similar written consent had been secured by the original investigator. Of course, it is the responsibility of the original investigator to consider the qualifications of any individual wishing to work with an existing data file, and it is particularly important to consider the individual's ethical record.

A related problem occurs when a study, begun as a one-shot study, is subsequently expanded to a longitudinal project. In such an instance, subjects' informed consent was only for the original data collection and not for the follow-ups since these were planned after the consent was obtained. The ethical procedure, then, would be to obtain a new consent agreement, reflecting respondents awareness of the expanded scope of the study.

This ethical question we have been discussing takes on a new light when looked at in the context of centralized data banks. Already, a number of such banks have been developed, for instance the Henry A. Murray Data Resource and Research Center at Radcliffe College. The Radcliffe Data Center was designed

in 1976 as a repository of data on educated women, with a particular emphasis on longitudinal research. Accordingly, staff members of that Center have set out to identify data sets which are relevant to their areas of interest, to serve as a clearinghouse for information about these data sets, to collect a number of the data sets and to sponsor research using these sets. Annotated listings of data resources are available from the Center. The ethical question which comes to mind in regard to such centralized data bank is: Doesn't the existence of centralized data bank threaten ethical conduct with regard to respondents' rights of informed consent and confidentiality? The question can be seen to be more complex here since the intermediary position of the data bank makes it more difficult for the original researcher to keep control over the rights of his or her respondents. Moreover, the subsequent researcher may feel less responsibility for the rights of these same respondents since the data were obtained from an organization and not from the subjects themselves, nor from the original investigator. The responsibility for the ethical conduct of subsequent researchers, thus, falls to the staff of the data banks. The Radcliffe Center has taken this responsibility quite seriously as is attested to by the contractual agreement routinely drawn up between the Center and any data contributors. This agreement spells out whether the original investigator wants: 1) anonymity of respondents maintained 2) certain types of researchers prohibited from having access to the data file 3) conditions set on whether and how follow-ups may be conducted and 4) copyright of the data collection instruments given to the Center. Furthermore, the Center warrants that it

> "will not violate any prior agreements between the contributor and subjects which are made known in writing to the Center. Contributor warrants that he/she has attached any agreement or consent forms or copies thereof which he/she or others obtained from the subjects of the contributed material, and to which he/she has access; that he/she has made reasonable efforts to obtain a sample of any agreement or of any consent form given to subject, in the study, and that he/she has on Exhibit A attached hereto described to the best of his/her recollection the information given to the subjects as to the future use of material, and the conditions under which the subjects agreed to participate in the studies. The Center will obtain an undertaking from users that they will not knowingly divulge any information which could be used to identify individual subjects in the dataset, except to the extent necessary for permitted follow-up studies or where the contributor indicates in writing that there is no reason for anonymity and the Center concurs in such conclusion." (Radcliffe Data Resource and Research Center, no date, p. 2)

The tremendous potential for abuse with a proliferation of centralized data bank exists. If these banks are to succeed while maintaining the integrity of social science research, it will be the responsibility of every researcher to do his or her part to behave in an ethical fashion. Moreover, contributors to data banks will have the added responsibility of ascertaining the reputability of these banks prior to providing data sets. This responsibility is not so different from that the original investigators face in deciding whether to allow other investigators to do secondary analyses on their data sets or to follow-up their respondents. It is more critical, however, as wrong choice of a data bank could have greater implications than misplaced confidence in a single investigator.

The Use of Deception and Unethical Research Practices

Whenever possible, the researcher should avoid the use of deception or of any measures which cause discomfort to subjects. The investigator is obligated to search for any reasonable way of designing a study which avoids deceiving subjects or causing them discomfort, even if the solution is difficult to find.

This ethical practice is supported by the Department of Health, Education and Welfare which has issued guidelines for research projects which involve human subjects. These guidelines caution against submitting individuals to research conditions which will expose them to physical, psychological, social and legal risks. Certain categories of individuals are more vulnerable to these risks. These include: fetuses, pregnant women, children, the mentally disabled, prisoners and other individuals "whose ability to give voluntary informed consent may be in question." The guidelines do not specify the kinds of research situations which may cause risks to the subjects. However, most of the risk categories are fairly obvious. Thus, it is clear that drugs of unproven effectiveness which could be damaging to the health of subjects would be considered a physical risk. However, what makes up a psychological risk? The use of deception could subject the respondent to psychological risk. Asking certain kinds of questions on questionnaires or in interviews could restimulate painful past occurrences for the individual, thus also creating psychological risk. The ethical issues surrounding both of these are discussed below.

Many research questions in the social sciences can be answered only by designing experiences which cause subjects some psychological discomfort, anxiety, embarrassment or annoyance. For example, it is difficult to investigate the effects of anxiety except in situations in which people experience anxiety. Of course, a researcher must weigh the benefits of experimentally causing anxiety in subjects in a laboratory situation as contrasted to studying the naturally occurring correlates of anxiety.

In addition, sometimes it is impossible to answer particular research questions without concealing the true nature of the study by the use of deceptive devices. This is because many research procedures invite bias due to defensiveness of respondents or their unwillingness to give anything but a socially desirable response. However, in deception, the investigator is not only misleading the respondents, but also invading their privacy by extracting data under false pretenses. Deception in research may occur in subtle forms as well: Any projective technique, for example, is a deceptive device unless a subject is forewarned of the true purpose of the instrument. This is true whether the researcher introduces the instrument under a different name (e.g., saying that the Thematic Apperception Test is a test of creativity rather than one which assesses intra-psychic processes) or does not describe it at all. Moreover, a simple interview often contains subtle elements of deceptiveness: The interviewer gains the confidence of the respondent by simulating agreement with his or her attitudes or by behaving in a pseudo-friendly manner which may seduce the respondent into "voluntarily" revealing inner thoughts or feelings under false pretenses.

Deception, even when innocuous, presents an ethical problem simply because a lie has been told. Moreover, whether the deception is innocuous or not, if

a subject has not agreed in advance to allow the researcher to observe a specific aspect of behavior, his or her privacy is being invaded. This is true whether or not the study in which one has agreed to participate is more or less pleasant than that one eventually experiences.

Often the deception is not innocuous, but results in some anguish, upset or discomfort on the part of the subject as in the now famous Milgram study. The Milgram study resulted in emotional trauma for many of his subjects, in spite of careful debriefing. When informed of the deception, many subjects were seriously disturbed by the knowledge that under certain circumstances they were willing to administer lethal levels of shock to another human being. The fact that the experiments involved deception and that no shock was actually being administered did not change respondents' reaction to their potential for causing harm to another human being.

Loss of dignity of the respondent may occur when no deception is involved. Even the most innocent-looking procedure may cause a respondent to feel uneasy or diminished. Since the researcher-subject relationship is one of unequal status, it is relatively easy for an investigator to make a respondent feel powerless.

Of particular potential harm are experimental practices, or interview or questionnaire questions which restimulate painful past experiences. For example, studies of victims of kidnapping may elicit from respondents a great deal of information about the experience. Thinking about the kidnapping, even for the purpose of responding to an anonymous questionnaire, could have the effect of causing the individual psychological trauma, even though the occurrence being invenstigated may be several years old. Thus, investigators must be sensitive to the psychological effect of every question asked of subjects.

Social Implications of Research

A final ethical area to be discussed relates to the social responsibility of researchers. It considers related questions: "Are there certain issues which should not be researched?" and: "Are there certain research findings which should be withheld at least temporarily for the benefit of humankind?"

While the ethical issues we have discussed in prior sections lent themselves uniquely well to solutions, in this case, the solutions are less clearcut, with more room for the individual conscience to manifest itself. Thus, in considering whether certain issues ought not to be researched or ought not to be revealed at this time, we run into moral injunctions which may conflict with the doctrine of free scientific inquiry. For example, recent reports of scientists' increasing success in directly manipulating genetic material has been met with moral outrage and concern, both from within and without the scientific community. While such genetic research techniques have clear scientific merit and may eventually save many lives and improve the lives of many others, their usage directly conflicts with prevailing religious and moral tenets about the creation of life. Similarly while many hailed recent successes in the fertilization of human eggs in vitro, the later implantation in the mothers and the subsequent normal deliveries and births, others were outraged and frightened at the potential dangers presented by this type of technology. The issue here has to do with the fact that current advances in technology outdistance religious and moral changes.

Other research topics which present ethical problems for investigators are any which, because of their moral or philosophical nature or their social implications, may be of particular interest to the lay public. Often, research findings are difficult for non-scientists, particularly members of the media to correctly interpret. It is not uncommon to find a newspaper article capitalizing on what the researcher calls "trends" in non-scientific findings. Nor is it unusual for a reporter to describe data based on correlational analyses as "proving" some outcome of interest. Large sensational headlines often ensue from such distortions, misleading the public as to what the research has actually demonstrated. Recall the furor over Jenson's work on intellectual differences between Blacks and Whites and imagine a typical headline: "BLACKS SHOWN TO BE GENETICALLY INFERIOR!" Given the respect with which scientists and researchers are held by the public, it would be understandable that readers of such a headline, and of the article which would accompany it, could conclude that Blacks have been conclusively demonstrated to be inferior to Whites. The controversy in the scientific community over Jensen's work probably did not reach the public, thus leaving some members of our society with ammunition to stoke their bigotry.

So, in addition to considering whether certain topics should be researched, investigators must carefully weigh whether and how certain research findings should be presented to the public. Since there are few definitive guidelines about where to draw the line, the individual researcher is left to decide whether the emphasis should be on free scientific inquiry above all else, on moral or philosophical injunctions alone or on some combination. Hopefully, in making this decision, researchers will find that they can sufficiently modify one set of values to be consonant with the other set of values and yet continue to produce scientific work of merit which will benefit humankind.

References

ACE Study on Campus Unrest: Questions for Behavioral Scientists, Science, 11 July, 1969, 165 (3889), 157-161.

Astin, A.W. & Boruch, R.F. A "link" system for assuring confidentiality of research data in longitudinal studies, American Educational Research Journal, 7(4), 1970, 59-67.

Privacy and Behavioral Research. Panel Report: Executive Office of the President, Office of Science and Technology, United States Government Printing Office, (February) 1967. Radcliffe Data Resource and Research Center, Memorandum of Agreement, no date. Russell Sage Foundation, Guidelines for the Collection, Maintenance and Dissemination of Pupil Records, Russell Sage Foundation Conference Report, 1970.

United States Department of Health, Education and Welfare, Public Health Service. Protection of the individual as a Research Subject--Grants, Awards, Contracts. Washington, D.C.: United States Government Printing Office, May 1969.

83033

Chapter 11

How the Ascertained Longitudinal Projects Might

be Utilized in Criminological Research

Sarnoff A. Mednick

William F. Gabrielli, Jr.

Katherine Teilmann Van Dusen

University of Southern California
Social Science Research Institute

How the Ascertained Longitudinal Projects Might
be Utilized in Criminological Research

Sarnoff A. Mednick, Ph.D., Dr. Med.
William F. Gabrielli, Ph.D
Katherine Teilmann Van Dusen, Ph.D.

University of Southern California
Social Science Research Institute

In this chapter we will attempt to specify how important questions in criminologic practice and theory may be answered by analyses of existing prospective, longitudinal studies. We will attempt to structure the material so as to maximize the usefulness of this presentation to criminologists interested in secondary data analyses using these longitudinal studies.

Unfortunately, it is impractical to hope that we will be able to explicate the importance of each of these studies to every possible question in criminology. The opportunities reported below can be seen as examples of the types of questions which are appropriate to longitudinal research and the variety of answers which are possible. We will not repeat the specific advantages of using these studies in understanding the causes of crime. The advantages have been amply described in this report.

One great benefit of using existing longitudinal studies rests with the fact that the populations are identified and examined before antisocial behavior becomes manifest. In other words these studies are most useful when they speak to etiological questions. We have, consequently, restricted our attention in this chapter to studies which first examined their subjects during infancy, childhood or early adolescence.

Most of these studies have multiple initial measures and may be utilized to answer different criminologic questions. We have organized the presentation around several criminologic questions. The first time a project is recommended it is briefly described. When it is mentioned again the reader is referred to the initial description. The descriptions are brief, containing, we hope, some useful information. We urge the interested reader to consult original publications or to examine the chapters each project has prepared for a volume which will soon appear (Mednick and Harway, in press).

The presentation is organized so that we proceed from less complex criminologic questions (biological) to more complex questions (neighborhood).

BIOLOGICAL AND COGNITIVE ANTECEDENTS OF CRIMINAL BEHAVIOR

Birth Damage and Violence

Violent crime is one of the critical problems in American Society today. A recent review article (Mednick, Pollock, Volavka, and Gabrielli, 1982) has reviewed evidence that part of the basis of violent criminal behavior is early damage to brain centers concerned with inhibition of behavior. One of the most important possible causes of such damage is perinatal difficulties. It is well known that the disadvantaged suffer a disproportionate amount of perinatal complications and also contribute disproportionately to violent crime. In addition, theories exist linking neurological dysfunction with impulsive violent criminal behavior.

One possible avenue for determining the relationship between perinatal brain damage and violent criminal behavior would be the examination of birth cohort data. Birth cohorts are frequently initiated in order to answer questions relating to perinatal factors and neurological disorder. Consequently these studies have especially careful and detailed coding of the pregnancy and birth process and frequently also include follow-up neurological examination. If we can identify individuals in the cohorts who have suffered early neurological damage, and also have been involved in violent crime, it would be possible to learn much about the role of brain damage in the etiology of violent crime. The remaining members of the birth cohort can serve as good controls.

The hypothesis was, in part, suggested by our experience with a Danish perinatal cohort. In Copenhagen a perinatal cohort of 9,006 consecutive deliveries was carefully studied by a small group of senior physicians in 1959-1961. To take advantage of these data for criminologic purposes we assessed the cohort's official delinquency records in 1978; they were then 17-19 years of age. Based on the neuropsychological and electroencephalographic literature we hypothesized that those members of the cohort who had developed into violent individuals would have suffered increased levels of perinatal damage and subsequent neurological symptoms.

We developed a scoring system for the perinatal and neurological information and these scores were related to the violent offenses of these individuals. We found a distinctive positive relationship between recidivistic violent delinquency and the birth variables of perinatal difficulty and neurological disorder.

There are three studies among the longitudinal projects summarized in this report which could prove valuable for exploring the neurological dysfunction hypothesis in the etiology of violent crime. The Philadelphia Collaborative Project, the California Child Health and Development Project, and The Florida Longitudinal Project.

The Philadelphia Collaborative Perinatal Project (Maracek) This project is part of the large U.S. Collaborative Project. The objective of the original study was to explore the origins of cerebral palsy and other neurological conditions. A population of 9,636 subjects has been studied.

The perinatal health status and condition of the child was very carefully monitored. Neurological examinations occurred periodically until the child was eight years of age. Measures of health status and intellectual growth were repeatedly taken. The population in 1981 is from 15 to 22 years of age and consists of both sexes. Recently Marvin Wolfgang has gathered data on the criminality of the cohort in the Philadelphia area.

In view of the fact that both the neurological and perinatal information for these subjects and the information relating to criminal violence exists on the same data tape, it is not an exceptionally difficult task to determine the correlations among these factors. One great advantage of the Philadelphia Collaborative Project is the fact that neurological examinations were taken neonatally, at one year of age and at seven years of age. The subsequent violent behavior in any of the individuals can be considered in the light of developmental patterns of neurological deficits.

The U.S. Collaborative Project in total included over 50,000 individuals at a number of cities across the U.S. An attempt is currently being made to collate the data into accessible magnetic tape form. In order to repeat the perinatal-violence analyses suggested above, it would only be necessary to ascertain the criminality of the rest of the Collaborative Project Cohorts.

California Child Health and Development Studies (Van Den Berg)
Along with the collaborative Project mentioned above, this study is one of the
world's major longitudinal birth cohort investigations. A total population
of over 20,726 births beginning with deliveries in 1959 was studied. At this
writing, the subjects are twenty-two years of age. The investigators have measures
of neurological functioning neonatally, at one year of age, five years of age and
at fifteen to seventeen years of age.

The investigators have not been interested in the problem of delinquency
or criminal behavior and no attempt has been made to assess these outcomes.
California, however, has one of the best criminal record systems in the U.S.
It should be possible to make use of these records in order to check
the entire population for criminal behavior. If we could determine which of the
individuals in this cohort has behaved violently it would be possible to conduct
an independent test of the hypothesis that early neurological damage is related
to later violent criminal behavior.

The Florida Longitudinal Project (Fletcher, Satz, Morris) The Florida
Longitudinal Project was originally designed to study the role of learning
disabilities in the development of serious emotional and behavioral disturbances.
Below we discuss the usefulness of this study in understanding the relationship
between learning disability and delinquency. The investigators examined a cohort
of all of the 497 white, male kindergarten pupils in Alachua County, Florida.
Measures collected on subjects included several indicants of neuropsychological
functioning at various time points during the school career. Much of the exist-
ing research on violent criminals which implicates brain damage is based on studies
utilizing neuropsychological testing of prison populations. As a result, questions
can be raised whether the brain damage preceded the violence (and thus could be
considered among the causes) or whether the brain damage was a result of the
individual's violent behavior. In the case of the Florida Longitudinal Study the
neuropsychological measures were taken at the kindergarten level and in the two
or three years following this and it is impossible that any brain damage dis-
closed by these neurological measures were caused by the later violent behavior
of some of these children.

In Florida it is possible to obtain criminal records and also delinquency
records from the official files. This could be done at minimal cost and would
provide a third replication of research investigating the relationship between
brain damage and violence.

In Chapter 5, Nesselrode and Baltes relate the importance of cross-cohort
replication in assessing the reliability of empirical relationships. The U.S.
cohorts come from decidedly different sections of the country; if we found in all
cohorts that individuals who evidence early neurological difficulties
of these cohorts that individuals who evidence early neurological difficulties
later are more likely to behave in a violent criminal manner, this would be very
strong evidence that early brain damage is a partial cause of criminal violent
behavior. One caution, however, there are a considerable number of studies in
the literature which demonstrate that biological variables such as perinatal
neurological status do not predict, in a simple manner, to later behavioral diffi-
culty. In a study by Drillien (1964), it was shown that premature infants later
have difficulty in learning. Additional analyses demonstrated that the relation-
ship was restricted to those premature children who grew up in an unstable
family setting. Premature children who grew up in a stable family evidenced no
important learning disability. We must also be very sensitive to such biosocial
interactions in the area of violence. It would be important to examine the data
from these three cohorts in order to evaluate the interaction of social factors
with possible early brain damage. The social and family data necessary for the

study of this interaction exist in these three cohorts.

Cognitive Factors and Delinquency

As Hirschi and Hindelang (1977) have pointed out, the role of the intellectual
and cognitive development in school performance and adjustment may be an important
building block in understanding the etiology of delinquency. A lively controversy
has developed in the literature around this issue. Evidence has been presented
both in support of, and against such a relationship. The several longitudinal
studies which have examined this question (West and Farrington, (1977) Wolfgang, Sellin
and Figlio (1972) and Kirkegaard-Sørensen and Mednick, 1977) all support the hypothesis
that pre-delinquents have lower tested intelligence, particularly verbal intelligence.
It would be of great value to examine carefully the many facets and details of this
cognitive ability--delinquency relationship. Fortunately, our longitudinal studies
yield an exceptionally rich source of information concerning intellectual perfor-
mance. The measures range from kindergarten to high school, from learning dis-
abilities, achievement test scores, school grades to teacher reports.

Reading readiness

Reading readiness is typically assessed at age 5 or 6 years, often in kinder-
garten. There are five studies among the longitudinal projects which examined
the reading readiness of their subjects.

Philadelphia High-Risk Study (Spivack) The investigators questioned whether
early school behavior would predict delinquency in a socially disadvantaged high
risk group. The investigators asked all 56 kindergarten teachers in the 29 schools
of the four center-city Philadelphia school districts to rate randomly selected
male and female school children from their classes. No teacher rated more than 12
children. The children were followed-up continually in later grades through high
school.

Among the data collected was a standardized test of reading readiness. Other
measures include responses to questionnaire items designed to indicate signs of
adjustment problems or other school problems, the Hahnemann High School Behavior
scale, the Teacher Rated Adjustment Scale (TRAS), "pink slip" information (school
records of activities meriting disciplinary action), counselor information about
school problems, student scholastic history and achievement data, and routine
psychological testing information. Police records on arrests are included in
the data bank. At the time of this writing (December, 1981) the subjects are
about 19 years of age.

Quincy Epidemiological Study of Behavior Problems. (Reinherz) The aim of
this investigation was to describe the emotional and behavioral development of 714
preschool male and female children in Quincy, Massachusetts in an attempt to pre-
dict outcomes. The efforts included a preschool screening battery giving infor-
mation on reading readiness, health status, demography, developmental status,
cognitive ability, sensory functioning, and behavior patterns. During the
kindergarten year, and grades 1, 2, and 3, school adjustment status, school per-
formance, behavior, family events and some personal adjustment measures were
taken. During the third grade evaluation, peer ratings, self-concepts and locus
of control measures were assessed. Developmental health, sensory functioning
and cognitive functioning were also again measured.

These children (in 1981) are nine years of age. They are at an ideal point
to begin a prospective, delinquency study.

Schooling of Young Children (Entwistle and Hayduk) The investigators described the social and psychological circumstances of young children's experiences in the first three grades. A total of 1,101 children were examined in their first grade in white middle class, white lower class and black lower class schools. Examinations of reading readiness, math ability, IQ, parents, childrens' and peers' estimates of childs abilities were included. Also assessed were classroom conduct, school absence, race, SES, and neighborhood quality.

The children (males and females) are (in 1981) 17-19 years of age; their delinquency has not been assessed.

Yesterday's Children. (Kraus) This study assessed the educational development of 274 New York City male and female children. The assessments began in first grade and continued through high school. The assessments included reading readiness, intellectual and achievement measures, personality evaluations, social questionnaires, parent attitudes and teacher judgements. The sheer extent of this rich data bank make it quite exciting despite the relatively small number of subjects. The subjects were selected from two schools, one was predominantly black, one predominatly white. The subjects are (in 1981) now 34 years of age. Delinquency information is not available in the data bank on these individuals.

High-Scope Longitudinal Preschool Evaluations. (Weikart and Schwenbart) This project examines a small number of (N=215) of low IQ (60-90) low SES children. They are very thoroughly examined with measures of reading readiness, scholastic aptitude, IQ, school achievement and behavior, social characteristics, mother's attitudes, records of hospitalization, school grades, and attitudes toward the school. In adolescence they investigated truancy, dropouts, disciplinary actions, employment activity, police contacts and court records.

Learning Disability.

There are many reasons that learning disability (LD) may prove an important antecedent to delinquency. First, in terms of Hirschi's version of control theory, learning-disabled individuals may find it hard to obtain rewards in school life and may not form attachments to that institution. In some cases, learning disability may be related to neurological disorder and may predict to later violent criminal behavior (as explained above). Learning disability is also related to attention disorders and hyperactivity which have been shown to relate to delinquency. In the following studies LD has been assessed early in the school career. Where antisocial behavior has been ascertained, this is noted.

There are three studies which have made specific efforts to assess LD. Outstanding among these is the (Fletcher, Satz, Morris) project.

Florida Longitudinal Project. This project was briefly described above in the section on neurological factors and violence. The project leader, Dr. Paul Satz, is one of the world authorities on LD. The test battery, repeatedly administered in 1970 to all of the 497 male, kindergarten children in Alachua County, Florida, yields a thorough picture of LD in this population. The boys are now (1981) 17 years of age. It would be possible to ascertain their antisocial behavior in Florida official police files.

Quincy Epidemiological Study of Behavior Problems. (Reinherz) This project is described in the Reading Readiness section; in addition to measures of reading readiness, the investigators made a complete assessment of LD

in their sample.

Woodlawn Mental Health Project (Kellam, Ensminger, Branch, Brown, Fleming) The Woodlawn Project is a study of children and their families from a black, urban, poor community on Chicago's south side. The total group consists of four successive cohorts of all first-grade children in the Woodlawn community. It started in 1963 and the children are now twenty-four years of age. They were studied in their first-grade classrooms and family data were collected through interview with the mothers. Data were gathered on learning disabilities, aggressiveness, acceptance of authority and social functioning; also teachers rated the child's behavior. At age 16-17, there was another assessment including measures of self esteem, social functioning, school records, and psychiatric symptoms. An extremely detailed (and innovative) description of family structure was developed; drug and alcohol usage statements and self-reports of delinquency were obtained. The project includes 1,252 children of both sexes.

School achievement.

This variable is assessed by standardized achievement tests and school grades. The measure is often made at several points during the individuals' school careers. These repeated measures could permit investigation of developmental changes in achievement. It would be of interest (in terms of intervention) to know whether those who eventually become delinquent begin their school career with poor achievement or if the achievement level drops in association with the beginning of a delinquent mode of life.

A very large number of the projects have repeated measures of school achievement. In most cases the measures are standardized school achievement examination. The same standardized tests are used in several different cohorts. The possibility of establishing cross-cohort reliability of these relationships by utilizing these cohorts are clear.

Florida Longitudinal Project (Fletcher, Satz, Morris) This project has been described above. Delinquency measures must be obtained.

Southern Youth Study (Picou, Thomas, Cosby). This project has examined factors in the adjustment of individuals and families in low-income farm areas, looking specifically at career development and decision-making. The studied youth were from several Southern rural and economically disadvantaged counties in Alabama, Georgia, and Mississippi. The investigators examined the educational, income, population, and occupational characteristics of the areas. They then related these factors to characteristics of the individual's situation (such as parents' SES, parents' occupation, race, sex, and religion) and to the individual's performance and values (school performance, interpersonal relations, occupational choice and encouragement, career opportunities and blockages, and work values). Outcome measures also included career opportunities, adult attainment including income, education, and occupation.

The data set provides an opportunity to examine a homogeneous population at high risk for delinquency because of social deprivation and frustrations due to imagined or real impediments to career objective. An advantage of this data set is the opportunity to examine the development of delinquency in a rural setting. Most of the projects in this review include urban populations. The project includes 7,972 subjects who were first contacted in 1966. There are four waves of data

collection. Delinquency has not been ascertained.

The Educational Testing Service Study (Hilton) The Educational Testing Service has examined longitudinal growth patterns among American youth. Fifth, seventh, ninth, and eleventh grade students were studied every two years until high school graduation (N=901). Data included test measures of school achievement; understanding, and intellectual skills. Included among the tests were the School and College Ability Test, Sequential Tests of Educational Programs and a test of general information. In grade 12 the students were administered the College Board American History Test, The College Board English Composition Test and the Preliminary Scholastic Aptitude Test. Information also available on the subjects includes background and experience (from a questionnaire), school grades, and information about fathers' education.

The oldest subjects are now about 30 years of age (1981). The population (N=45,901) is representative of the U.S., and is enormous. A federal source of crime registration might be useful here since there has been no delinquency ascertainment.

Philadelphia High Risk Study (Spivack) This study is described above in the section on Reading Readiness.

Project Talent (Steel, Wise Abeles) The investigators have examined the personal, educational, and experiental factors which promote or inhibit the development of human talents. The investigation involved testing of over 375,000 American high school students in 1960 from 987 high schools. Data collected included results of achievement tests, language aptitude and ability tests, complex intellectual aptitude tests, visualization tests, mathematics tests, and clerical and perceptual aptitude tests. Follow-up information included one-year, five-year, and eleven-year educational experience, career development, and personal information (e.g., race, religion, martial status, health).

Potential advantages of this data set are its representativeness and completeness. The data could be used for extensive investigation of educational development patterns which can relate to the etiology of delinquent and criminal behavior. Maturational patterns can be related to the onset and development of officially recorded criminal behavior. The completeness of the aptitude measures may be useful in determining the profile of aptitude type which is most amenable to criminological influences and the types which are least vulnerable.

The youngest subjects (males and females) are now (in 1981) 36 years of age. Criminological information is not available in the data bank.

Adolescence to Young Adulthood: A Twelve Year Prospective Study (Jessor and Jessor) This project describes the course of psychosocial development over time in order to predict the onset of certain developmental transition behaviors, and to explore the continuity between early and later developmental states. They studied 1126 seventh, eighth and ninth grade students of a junior high school, and 462 freshmen of a college of arts and sciences. Data consist of responses to items of a 50-page questionnare recording details of personality, social interaction, behavior, and demographic characteristics. School achievement data were also obtained.

At the time of the initial assessment the subjects (both sexes) ranged in age from 13-19 years. They are currently (1981) 25-31 years of age. Perhaps the most unique and useful aspect of the study in relation to criminology is the concentration on measures of personality, personal belief structure, motivations,

internal control sturctures and parent and friend support and controls. Self report/measures of antisocial behavior as well as drug use are included.

Youth in Transition Project (Bachman and O'Malley) These investigators have attempted to relate social environments and experiences to school performance and occupational adjustment. They studied 2,213 tenth grade boys in 87 Michigan public high schools. The subjects were examined five times over an 8-year period. Variables included measures of family stability and background, ability and educational attainment. Measures of job satisfaction, self-esteem and school misbehavior and drug abuse were included. The subjects (all males) are now (1981) 30 years of age. With ascertainment of adult, criminal behavior this project has important potentials.

School and Family Effects on Student Development (Epstein) This project includes 14,604 fourth grade children (both sexes) from small city, suburban and rural communities in Maryland. The subjects are now (1981) 19-25 years of age. The intent of the study is to examine how family and school rules and the possibility to enter into decision-making influence the child's school achievement. Extensive assessment of achievement, IQ, teacher ratings, family processes, race, family structure, type of school, friendships, sociometry, and school grades are included.

High-Scope Longitudinal Pre-School Evaluation (Weikart and Schweinhart) This project is described above in the section on Reading Readiness

Yesterday's Children (Kraus) This project is described above in the section on Reading Readiness.

Intelligence test scores

The relationship of tested intelligence to delinquency has aroused a lively controversy in the literature. An important question in evaluating the etiological significance of this relationship is the extent to which lower test intelligence of the delinquent may result from his estrangement from academic pursuits (the importance of which has been indicated by most delinquency theories). The issue can be best studied by examination of early IQ test scores of future delinquents. Many of the longitudinal studies included IQ measures in their assessments.

Psychosocial Development of Aggressive Behavior (Eron, Huessman, Walder, Lefkowitz) This is a longitudinal study of the development of aggression. The subjects are all third-grade school children in Columbia County, New York in 1960 (N=875). The data represent children in a rural New York county. The investigators hypothesized that "aggression is learned by a child from interactions with the environment." Original data included IQ tests, family interviews on discipline and punishment, student rating of aggression in other students and interview information from parents. The children were again studied at age 19 with aggression scales, the MMPI and interviews.

Since the 1960 data include information about the child's home environment (including the amount of punishment and the opportunity for witnessing aggression in the parents), and since bonds (or attachments, or identification) with important others was assessed, many important hypotheses on the role of the family in the development of violent behavior could be examined. The subjects (both sexes) are now (1981) 30 years of age. They reside in a low-migration area of the U.S. Official crime records should be available.

Florida Longitudinal Study (Fletcher, Satz, Morris) This study was described in the section on Birth Factors and Violence.

Educational Testing Service Study (Hilton) This study was described in the section on School Achievement.

Philadelphia High Risk Study (Spivack) This study was described in the section on Reading Readiness.

Project Talent (Steel, Wise, Abeles) This study was described in the section on School Achievement.

Youth in Transition Project (Bachman and O'Malley) This project was described in the section on School Achievement.

Quincy Epidemiological Study of Behavior Problems (Reinherz) This study was described in the section on Reading Readiness.

School Family Effects on Student Development (Epstein) This study is described in the section on School Achievement.

Yesterday's Children (Kraus) This project is described in the section on Reading Readiness.

PERSONALITY

Certain personality characteristics have face validity as predictors of delinquency. These include lack of ability to delay gratification, impulsiveness, rebelliousness, and aggressiveness. Other personality characteristics may be seen as related to theories of delinquency causation. These include level of aspiration (strain theory) self image and self esteem (labelling and some strain theories) and alienation (strain theory). Some of the longitudinal projects ascertained have childhood and youth measures of these personality characteristics.

Inhibition of impulse

Two of the studies have very early measures of the child's ability to delay or temporarily inhibit gratification. One project (Block and Block) uses behavioral measures, the second project employs teachers' ratings (Spivack). The Spivack project includes measures of delinquency

Personality and Cognitive Development (Block and Block) The aim of this project was to investigate the development of ego-control and ego-resiliency. The investigators chose to examine these factors in preschoolers (age three) to avoid the effects of school. They report on follow-up examinations at ages four, five, and seven.

The children are 130 males and females drawn from two nursery schools in Berkeley, California. They measured delay of gratification, persistence in overcoming barriers in order to get to goals, motor inhibition, level of aspiration, distractability, and planning orientation. The children are now (1981) 15 years of age. This project is mentioned here despite the small sample because of the potential special relevance of the measures to delinquency and the fact that these variables are rarely measured in longitudinal investigations.

Philadelphia High Risk Project (Spivack) The project is described in the section on Reading Readiness however, there is a variable of special interest to this section. A teacher rating of impulsiveness is included in this data

bank.

Self esteem

Measures of self esteem are quite common in these projects. They all use paper and pencil questionnaires administered in school. The following projects contain early measures of self esteem:

Youth in Transition (Bachman and O'Malley) This project is described in the School Achievement Section.

Health Consequences of Drug Use (Brunswick) This project involved the collection of data on drug use, health, development, disability, psychological well-being, self esteem, family characteristics, alienation, social networks, employment, mobility and educational status on 536 Central Harlem male and female youths, age 12-17. There was a second assessment in 1974. There was no delinquent behavior recorded (except for the drug use).

The drug users in this sample are at high risk for criminal behavior. They are now (1981) 24-29 years of age. If their criminal and delinquency records could be obtained the relationship of the drug use to the antisocial behavior would be interesting to study.

Adolescence to Young Adulthood (Jessor and Jessor) This project was described in the section on School Achievement.

Family and Peer Processes in Adolescent Drug Use Study (Kandel) The investigators examined the nature of family contexts and family behavior in which drug use develops. To this end, four data collections were conducted. The subjects are New York State secondary school students (two groups; N1=8,206, N2=7,250) with broad ethnic and racial sampling. The investigators (in 1971 and 1972) collected questionnaire data mailed to parents and questionnaires completed by students during school time. Efforts also included collection of the same information from all subjects absent from the schools on the day of administration of the original questionnaire. One sub-group (the 1972 senior class) also participated in a mail and telephone follow-up survey of the class five to nine months after graduation.

Self reports of delinquent behavior were collected.

School and Family Effects on Student Development (Epstein) This project is described in the section on School Achievement.

Woodlawn Mental Health Project (Kellam, Ensminger, Branch, Brown, Fleming) This project is described in the section on Learning Disability.

High-Scope Longitudinal Pre-school Evaluations (Weikart, Schweinhart) This project is described in the section on Reading Readiness.

Alienation

Alienation is a concept central to several sociological theories, most notably strain theories. Several versions of this area of theory postulate the societal blockage of aspiration attainment, which in turn is expected to lead to

alienation from the conventional social order. With such alienation should come freedom to violate norms, or even a sense of rebellion, leading to antisocial behavior.

Alienation is assessed in two projects by questionnaire.

Health Conseuqences of Druge Use (Brunswick)   This project is described in the section on Self Esteem.

Adolescence to Young Adulthood: A Twelve-year Prospective Study (Jessor and Jessor)   This project is described in the section on School Achievement.

Rebelliousness

While criminological theory does not usually concern itself with psychological variables, certain theories (e.g., strain theory) could be interpreted to predict that children who are more rebellious would tend later to be delinquent. A number of studies have assessed the rebelliousness of the school child. These measures include teacher ratings and self-report instruments.

Youth in Transition Project (Bachman and O'Malley)   This project is described in the section on School Achievement.

Woodlawn Mental Health Project (Kellam, Ensminger, Branch, Brown, Fleming) This project is described in the section on Learning Disability.

Philadelphia High Risk Study (Spivack) This project is described in the section on Reading Readiness.

Family and Peer Processed in Adolescent Druge Use Study (Kandel)   This project is described in the section on Inhibition of Impulse.

Aggression

Aggression in nursery, kindergarten and elementary school years has been found to be related to later delinquency (West and Farrington). It has also been shown to be a reliable personal characteristic. Kindergarten boys who are aggressive tend to be aggressive in adolescence (Olweus). Aggressive attitudes and behaviors in school years are readily observable. If this early characteristic is a reliable predictor of later delinquency, its use in prevention could easily be envisaged. Two projects assessed predelinquent aggressiveness. The first project (Eron et. al) was centered on aggression and contains multiple measures. The Bachman and O'Malley project has a questionnaire measure of tendency toward aggressiveness.

Psychosocial Development of Aggressive Behavior (Eron, Huessmann, Walder, Lefkowitz)   This project is described in the section on Intelligence Test Scores.

Youth in Transition Project (Bachman and O'Malley)   This project is described in the section on School Achievement.

Level of aspiration

As mentioned, an individual's level of aspiration and its thwarting or fulfillment is central to some theories of the etiology of delinquency (especially strain theories). Five of the projects have careful measures of this variable.

Panel Study of Social Stress (Menaghan)   This project is described in the section on Mental Illness and Mental Health.

Personality and Cognitive Development (Block and Block)   This project is described in the section on Inhibition of Impulse.

Adolescence to Young Adulthood: A Twelve-Year Prospective Study (Jessor and Jessor)   This project is described in the section on School Achievement.

Southern Youth Study (Picou, Thomas, Cosby)   This project is described in the section on School Achievement.

Youth in Transition Project (Bachman and O'Malley)   This project is described in the section on School Achievement.

High-Sco e Longitudinal Preschool Evaluations (Weikart, Schweinart)   This project is described in the section on Reading Readiness.

Sociometry

One of the longitudinal projects included sociometric assessments; (The School and Family Effects on Student Development Project, Epstein).

Mental Illness and Mental Health

There are a number of reasons it might be important to examine the relationship between mental illness and crime in a family. First we have the newly discovered problem of trans-institutionalization (Guttridge, 1981). There seems to be a tendency for great overlap of clients between the mental health and criminal justice systems.

In addition some research has shown higher levels of crime among the seriously mentally ill and among the children of parents who are seriously mentally ill (Kirkegaard-Sørensen and Mednick, 1977).

The following studies may serve as a basis for such an investigation:

Midtown Manhattan Longitudinal Study (Srole and Fischer) The Midtown Manhattan Study was launched in 1954 to describe the mental health of an urban population. The project began with an area probability sample of 1,660 adults ranging in age between 20 and 59 years. The sample was interviewed in their homes concerning their symtomatology. Measures of degrees of tension, anxiety, depression, tendencies to withdraw, paranoid ideation, excessive alcohol and food intake and sociopathic orientations were obtained. The information was almost completely geared to assessing mental health.

In 1974 the population was approached again. A total of 858 respondents were located alive (266 were certified as dead and 536 were unlocatable). Investigators found the 858 survivors to be representative of the total cohort. Of the 858 located, 695 consented to be interviewed.

Many of the 1954 items were used again in 1974 so as to maximize comparability. The general finding of the research was that mental health did not deteriorate over the 20 years as a function of living in New York City. The mental health of the women seemed to have improved over this period of time.

The subjects are now 47-86 years of age. The comparison of the mental health information with cumulative police records could be revealing.

City Children (Langner)   This project is described in the section on Family Discipline.

Panel Study of Social Stress (Menaghan)   The project focuses on the life problems, social contexts and coping responses of "ordinary people." The investigators are especially concerned with occupational, economic, marital and parental stress in adult life. They have completed two waves of interviews. In 1972, interviews were conducted with 2,300 people representative of the adult population of urbanized Chicago. The interviews assessed the wide range of problems and hardships people experience as workers, bread winners, husbands, wives and as parents. The interview also identifies the resources and responses they bring to bear in coping with these life strains; finally the interview enumerates the symptoms indicative of emotional stress and psychological disturbance.

In 1976, the investigators conducted a follow-up interview concerned with problems and challenges that converge on the lives of people. Interviewees varied in age between 18 and 65 years of age.

The project has certain benefits for study of criminological issues. First, it seems to be an ideal project for evaluation of the role of occupational dis-satisfaction in criminal behavior. The project has careful recording of the re-wards that the job holds and the degree of pressure and the noxiousness of the physical work associated with the job.

The project also has paid a good deal of attention to the parental life of the adults. In view of the careful attention paid to the parenting role it will be quite interesting to see how the children's behavior evolves as a function of these parental behaviors. It would also be interesting to examine the relation-ship of unemployment to criminality in this population.

It is necessary to obtain police records in order to utlize this project.

FAMILY FACTORS AND DELINQUENCY

The study of family factors as they relate to delinquency has a long tradition. Virtually no one disagrees with the notion that family characteristics and inter-actions play a role in the development of delinquent behavior. In most of the studies that appear in the literature, measures of family variables were taken at the same time (or after) measures of delinquency were taken. This places the causal sequence in doubt. The studies included in this report, then, would make a significant contribution to the literature by virtue of their prospective designs.

Additionally, while family factors are universally acknowledged as relevant to delinquent behavior, the absolute strengths of the relations with specific variables and their relative importances are far from established. Finally, more work should be done on the interactions of other variables with family factors. Fortunately, the studies surveyed here are a rich source of family data as well as potential interacting factors.

Family structure. Family composition has, perhaps, the longest history of

inclusion in delinquency studies. A related aspect of family conditions is stability, including number and types of changes. Such factors can be related to discipline, morality learning, and inhibition learning, among others. Certainly family structure and its changes can affect the attachments so central to control theory. These types of variables can be found in the following studies from this report:

The Houston Study (Kaplan)   This study is described in the section on Attachment to Schools.

Youth in Transition Project (Bachman, and O'Malley)   This study is described in the section on School Achievement.

Family and Peer Variables in Adolescent Drug Use History (Kandel)   This project is described in the section on Intelligence Test Scores.

Attachment to family.

This variable is central to several theories of delinquency, especially control theory. Different theories assign different roles to attachment in the etiology of delinquency, some contradictory to others. Some of these studies could provide critical tests of the role of this variable in the development of delinquent behavior.

Psychosocial Development of Aggressive Behavior (Eron, Huessmann, Walder, Lefkowitz)   This project is described in the section on Intelligence Test Scores.

Adolescence to Young Adulthood: A Twelve-Year Prospective Study (Jessor & Jessor)   This project is described in the section on School Achievement.

Family and Peer Processes in Adolescent Drug Use Study (Kandel)   This project is described in the section on Inhibition of Impulse.

Woodlawn Mental Health Project (Kellam, Ensminger, Branch, Brown, Fleming)   This project is described in the section on Learning Disability.

High-Scope Longitudinal, Pre-School Evaluations (Weikart, Schweinhart)   This project is described in the section on Reading Readiness.

Socioeconomic status

This variable is the pivotal ingredient in strain and some cultural deviance theories. There has, however, recently been a strong controversy over its impor-tance. The controversy was stimulated by the emergence of self-report studies of delinquent behavior. The controversy is far from settled and would profit by analyses of data in this group of studies, especially those that include self-report.

Psychosocial Development of Aggressive Behavior (Eron, Huessmann, Walder, Lfkowitz)   This project is described in the section on Intelligence Test Scores.

Florida Longitudinal Project (Fletcher, Satz, Morris)   This project is described in the section on Reading Readiness.

Southern Youth Study (Picou, Thomas, Cosby)   This project is described in

the section on School Achievement.

Educational Testing Service Study (Hilton)   This project is described in the section on Learning Disability.

Adolescence to Young Adulthood: A Twelve-Year Prospective Study (Jessor and Jessor)   This project is described in the section on School Achievement.

Family and Peer Processes in Adolescent Drug Use Study (Kandel)   This project is described in the section on Self Esteem.

Youth in Transition (Bachman and O'Malley)   This project is described in the section on School Achievement.

High-Scope Longitudinal Preschool Evaluations (Weikart, Schweinhart)   This project is described in the section on Reading Readiness.

Family Discipline

The role of disciplinary practices has been addressed in some past delinquency studies, but has not been central to most of them in spite of its obvious potential in predicting problem behavior through the learning of inhibition of antisocial behavior. More studies of this variable, in combination with others, are needed. The following studies from this group would allow these analyses:

Psychosocial Development of Aggressive Behavior (Eron, Huessmann, Walder, Lefkowitz). This project is described in the section on Intelligence Test Scores.

Adolescence to Young Adulthood A Twelve-Year Prospective Study (Jessor and Jessor). This project is described in the section on School Achievement.

Family and Peer Processes in Adolescent Drug Use Study (Kandel). This project is described in the section on Self Esteem.

City Children (Langner). The purpose of this project was to obtain information at the time a child was growing up to see what meaning the family environment had for his later mental health. The emphasis was placed upon family and the broad social environment with briefer attention given to physical illness, accidents and handicaps. A brief family history of emotional problems and their treatments was collected. Global measures of parental disturbance, education, marital and police history was obtained. It was also planned at a general epidemiological survey of mental health in a child population.

The population studied resided in the area between Houston Street and 125th Street on the East and West sides of Manhattan in New York City. A representative sample was taken of 1,034 families of which 56% were white, 29% were Spanish-speaking, 14% black and 1% other. The age of the children at the time of the interview was between 6 and 18 years.

A sample of individuals on the welfare rolls was taken from the same area; the sample consisted of 1,000 families (27% white, 36% Spanish-speaking, and 37% black).

This project has important implications for criminological research. The populations are representative of welfare patients and individuals in large urban areas. There is an individual evaluation of family size and structure and

family discipline. These can be related to delinquency and criminality outcomes which have been assessed. The sample is also interesting because it monitors not only delinquency but mental health. Langner expresses the prevalence of problems including mental health problems and delinquency problems; he also reports separately on these two problems.

The City Children Project can also be of great use in examining the effect of mental illness in the parents on delinquency and criminality in the children.

This project which is known mainly in psychiatric research circles should be given wider attention by individuals interested in research in antisocial behavior. Both sexes were studied; the children are now (1981) 21-38 years of age. As mentioned above, police arrest records were obtained.

Mobility

Frequent geographic moves by the family of a child can contribute to problem behavior in clear ways. For instance, many moves can break down bonds to school peers, thus leaving the child vulnerable to antisocial impulses. Very few studies have investigated this as a family variable, largely because of the difficulty in measuring it over time. Most studies of mobility have been ecological in nature. The following study can help to fill this gap.

Self Attitudes and Deviant Responses (Kaplan). This project is described in the section on School Attachment.

ATTACHMENT TO SCHOOLS

Hirschi (1969) has suggested an important theory concerning the causes of delinquency. A critical element of this theory asserts that children who do not develop an important attachment to the institution of school have a higher likelihood of later developing delinquent behavior. In Hirschi's own study he demonstrated that children who were delinquent evidenced poor attachment to school. Support for the theory would be strengthened considerably if it could be shown that the poor attachment to school preceded the development of delinquent behavior. Such a study has not been accomplished; however, it would be a relatively simple matter to examine this question within the context of existing longitudinal research.

We have located five projects which have detailed data concerning children's early attitudes toward the schools and their teachers. It happens that, in the case of four of these projects, data were also gathered on the children's delinquent acts. In one case the information consists of police records, and in three other cases, of self-report. We will briefly describe each of these studies and its relationship to the research question.

The Philadelphia High-Risk Study (Spivack). In this study the investigator is interested in learning if early school behavior will predict delinquency in a socially disadvantaged group. A total of 658 children who were in kindergarten in 1968 were randomly selected. The teacher rated the school behavior of each child including items indicative of the child's relationship to school and the teacher. The records of criminal behavior include the delinquency the child exhibited at the schools and also reflected in police records. The project is also described in the section on Reading Readiness.

Family and Peer Processes in Adolescent Drug Use Study (Kandel). The investigators in this study are attempting to examine the relative influence of family

and peers on adolescent drug use. The study involved two groups of New York secondary school students (one group totalling 8,206 children, the second group totalling 7,250 children); the data are based on questionnaires mailed to parents and questionnaires completed by students during school time. The usual problem of delinquent children being those absent from the school the day of questionnaire administration was attacked by special efforts to include data from these subjects. There is a group of questions relating to the child's attachment to the school; the investigators also gathered information on the child's delinquent involvement. This project is also described in the section on Self Esteem.

The Houston Study (Kaplan) This project studied all of the seventh-grade students in half of the thirty-six junior high schools in the Houston independent school district. The efforts were aimed at testing the theory that self attitudes influence the emergence of deviant behavior. To this end they studied the students for two years, testing them three times during this period. The data include measures of attitudes toward the schools and the teachers. Delinquency was assessed by self-report. It would be interesting to learn if seventh-grade students' attitudes toward school would predict to their later self-reports of delinquency.

School and Family Effects on Student Development (Epstein) In this project conducted in small-city, surburban, and rural communities in Maryland, fourth-grade children were followed for some years and retested. The aim of the study was to explore the effect on school achievement of participation in school-and family-decision-making by children. As part of the assessment the questionnaire surveys the children's attitudes toward the schools and teachers in a variety of contexts. It would be possible to construct a scale measuring attachment to school and teachers. Information exists already on the amount of trouble that the students had in the classrooms. Separate items are available on playing hooky, skipping classes and on disciplinary suspension from schools. In addition the extent to which the student observed classroom deportment was estimated. Because of the strong relationship between delinquency and school misbehavior it would be possible to estimate the role of early school attachment to later antisocial behavior in the school. In view of the fact that these children all resided in one state (Maryland), it is conceivable that later delinquent behavior could be assessed from official records. The number of subjects in the study totals 14,604. This project is also described in the section on School Achievement.

Youth in Transition Project (Bachman and O'Malley) At the University of Michigan a group of investigators have attempted to relate social environments and experiences to school performance and work. They studied 2,213 tenth-grade boys. They were assessed five times over an eight-year period. Among the variables are measures which could be related to attachment to school. Among the outcome measures are self-report data on delinquency and drug abuse. This project is also described in the section on School Achievement.

INFLUENCE OF PEERS

It is well established that delinquency tends to occur as a group phenomenon. The question naturally arises as to the influence of delinquent peers on the development of delinquent behavior in an individual youngster. There are conflicting views on this question. Traditionally, and according to various forms cultural deviance theories, the child who has delinquent friends is more likely to become delinquent. According to Hirschi's expositions of control theory, however,

attachment to peers, regardless of their delinquent status insulates against delinquency. Others have found that it is the relative attachment to peers compared to parents that is important. Clearly, studies that include data on peers and subjects' attachment to them would be valuable in adding to this literature. The following studies hold such potential:

Family and Peer Processes in Adolescent Drug Use Study (Kandel) This project is described in the section on Self Esteem.

Youth in Transition Project (Bachman, and O'Malley) This project is described in the section on School Achievement.

Quincy Epidemiological Study of Behavior Problems (Reinherz) This project is described in the section on Reading Readiness.

School and Family Effects on Student Development (Epstein) This project is described in the section on Intelligence Test Scores.

Schooling of Young Children (Entwistle, Hayduk) This project is described in the section on Reading Readiness.

NEIGHBORHOOD

Both cultural deviance and some (ecological) control theorists stress the importance of neighborhood characteristics in generating delinquency. Neighborhood SES can have an impact quite separate from individual SES. Other factors are ethnic heterogeneity, mobility rates and crime rates. While these types of factors are not good predictors of individual delinquency, (they are better in predicting delinquency rates) they have proved valuable in specifying other relations such as those between family characteristics and delinquency and between biological factors and delinquency. These contextual effects have not been well researched, so those studies that include them would be quite fruitfully exploited.

Youth in Transition Studies (Bachman and O'Malley) This project is described in the section on School Achievement.

High-Scope Longitudinal Pre School Evaluations (Weikart, Schweinhart) This project is described in the section on Reading Readiness.

Schooling of Young Children (Entwistle, Hayduk) This project is described in the section on Reading Readiness.

COMMENTS AND SUGGESTIONS

In this chapter we have attempted to suggest how specific questions concerning the etiology of delinquency and crime may be attacked by utilization of existing longitudinal research projects. The emphasis on etiology caused us to concentrate on the presentation of studies encompassing the perinatal to adolescent period of development. Longitudinal studies, however, can answer many other types of questions of importance to the field of criminology.

For example, the specific deterrent effect of police action or court sanctions can be best evaluated by consideration of such long-term consequences as reduction to recidivism, or in seriousness of crimes or "density" of criminal activities. Questions regarding the differential etiology of different types of adult criminal behavior may be best studied through research which examines a cohort early in life and ascertains their criminal behavior in adulthood. Some

of the subjects in the studies described above are old enough now to make such an analysis possible.

Recommendations

It seems clear that a critical goal of criminological research must be the development of methods of prevention of criminal behavior. There are many possible paths to this goal. One important avenue to consider is intervention with the chronic offender (the 6-10% most offensive criminals). This small group of offenders accounts for most of the serious crimes. If we could devise techniques to substantially reduce the criminal activity of the chronic offender the resultant reduction in the total number of serious crimes might make it possible for us to be less punitive with the other 90% of offenders.

The critical first step in achieving this goal is to develop methods of early identification of the chronic offender. It would be extremely useful if the future chronic offender could be identified during the course of his early arrests. Longitudinal, prospective studies offer an improtant approach in the development of methods of early detection. (In Fact, prospective research may be the only possible approach).

Identifying the chronic offender. If we were able, early in criminal careers to identify successfully the youth who might become a chronic offender, the characteristics and experiences which distinguish him might teach us how to help him inhibit his extreme antisocial behavior. Even a modest rate of reduction of the criminal activity of the chronic offender would have a marked effect on the total level of crime. The potential enormous payoff to society, to the other 90% of offenders and to potential victims clearly warrants a major investment of societal resources in this research endeavor.

Is early identification of the chronic offender a practical possibility? Research on genetic, biological, family and social factors suggests that the chronic offender can be discriminated from other offenders (e.g., Mednick et. al, in press).

Earlier research has typically not been multidisciplinary; it has considered biological, psychological or social data in prediction. What is needed is a prospective study which combines these predictors and considers their interactions. Such a project will be a major undertaking; it must be prepared carefully.

A critical research decision in such a prospective study will be the choice of initial measures. These measures will form the basis of any selection procedure which will be developed. We would suggest that this choice be informed by the data of existing longitudinal research. Specifically, we propose that focussed, small teams of two-four longitudinal researchers be invited to come to a Center with their data tapes. These researchers and their projects whould be extremely carefully selected as to make possible definitive cross-cohort-replicated answers to specific early identification questions. For example, we might invite several investigators who have extensive school data and who also have identified the chronic offenders in their cohorts. (In some cases where the school data are superb and the criminal data are lacking, police and court records may be available for assessment of criminal behavior). In the proper setting, with proper assistance, intensive collaborative efforts would rather quickly delineate reliable school characteristics of the chronic offender. The power of cross-cohort-replications would reduce the analysis problems and hasten the research progress considerably.

The literature and our experience in working in longitudinal research suggests that this consortium effort will make it possible to develop a universe of extremely

promising, highly reliable, definitive factors which, in cross-cohort-replication, have proven useful in prediction of the chronic offender. These factors would then be examined in a cohort of young offenders to assess the combined ability of these measures to identify with minimal error those who will later become chronic offenders.

Our reading of the longitudinal studies and our extensive contacts with the researchers have convinced us that this plan is highly feasible and should be assigned high priority.

References

Mednick, S. A. and Harway, M., Longitudinal Research in the United States. Hinghman, Mass, Kluwer Nijhoff Publishing Company, in press.

Mednick, S.A., Pollock, V., Volavka, J., and Gabrielli, W. Biology and Violence. In M. Wolfgang and N. Weiner (Eds.), Proceedings on Criminal Violence, Sage Publications, in press.

Drillien, C. M. The growth and development of the prematurely born infant. Edinburgh: Livingston, 1964.

Hirschi, T. and Hindelang, M. J. Intelligence and delinquency: A revisionist review. American Sociological Review, 1977, 42, 571-587.

West D., and Farrington, D. P. The delinquent way of life. London: Heinemann, 1977

Wolfgang, M., Figlio R. M., and Sellin, T. Delinquency in a birth cohort. Chicago: University of Chicago Press, 1972.

Kirkegaard-Sørensen, L., and Mednick, S. A. A prospective study of predictors of criminality: 4. School behavior. In (Eds.) S. A. Mednick and K. O. Christiansen Biosocial bases of criminal behavior. New York: Gardner Press, 1977.

Olweus, D. Prediction of aggression. In (Eds.) S. A. Mednick and M. Harway, Longitudinal Research in the United States. Hinghman, Mass, Kluwer Nijhoff Publishing Company, in press.

Guttridge, Patricia. Psychiatric and non-psychiatric factors in the hospitalization of adolescents. T. S., 1981, unpublished doctoral dissertation, University of Southern California.

Kirkegarrd-Sørensen L. and Mednick, S. A. A prospective study of predictors of criminality: 2. A description of registered criminality in the high-risk and low-risk families. In (Eds.) S. A. Mednick and K. O. Christiansen Biosocial bases of criminal behavior. New York: Gardner Press, 1977.

Hirschi, T. Causes of delinquency. Berkeley: University of California Press, 1969.

END