PROGRAM EVALUATION IN CONTEXT: MODELS OF NEGOTIATED KNOWLEDGE PRODUCTION AND USE

by

Alberta Potter Levitan

B.A. Antioch College 1959
M.A. University of Michigan 1961

A Dissertation Submitted in Partial Fulfilment of the Requirements for the Degree of Doctor in Philosophy in

The Faculty of Graduate Studies (Department of Sociology)

We accept this dissertation as conforming to the required standard

The University of British Columbia
February 1983

© Alberta Potter Levitan, 1983
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Sociology

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date May 1, 1983
ABSTRACT

Recent growth and co-optation of social reform programs into the structure of the State, and parallel development of public policymaking, have precipitated closer linkages between social research and the policy intervention process. Program evaluation refers to a variety of descriptive and analytical studies of program process and/or program impact. Two models of program evaluation research are relied upon in design and implementation of evaluation studies: (1) a conventional model derived from a positivist paradigm of social research and (2) an alternative model evolved from an interpretive paradigm. Critical review of these models suggests their complementarity for comprehensive evaluation studies, but emphasizes the extent to which they minimize the significance of larger political/economic context of program development in shaping evaluation processes.

The purpose of this dissertation has been to develop a wider selection of evaluation research models which specifically take into account construction of the research "product" and characteristics of the larger structural context in which such products are designed to be used. The theoretical strategy relies on aspects of Strauss' negotiated order paradigm and approaches to policy research taken by Rein and Wiseman, and involves an effort to relate more stable structural characteristics of the social, political, economic and organizational context of reform programs to a series of six basic areas of negotiation in the evaluation process. These include: (1) delineation of major actors; (2) organizational placement of program evaluation work; (3) choice of general research strategy; (4) selection of appropriate research model and methodology; (5) construction and content of research reports; and (6) planning for research utilization.
This framework provided the theoretical perspective for description and analysis of four case studies: two in housing policy, one in private social service delivery and one in delivery of legal education services.

Conclusions from case studies, and other research suggested four models of negotiated knowledge. New models include elements of positivist and interpretive models but are designed around the structural context of program planning and implementation and focus directly on the six basic areas of negotiation. This expanded repertoire of models of negotiated knowledge production and use have been labelled Experimental, Managerial, Collaborative, and Transformative.
TABLE OF CONTENTS

List of Figures ..................................................... vii
Acknowledgments .................................................... viii
Preface ................................................................ ix

CHAPTER 1 SOCIAL REFORM, POLICY RESEARCH AND THE PROGRAM EVALUATION PROCESS ..................................................... 1

  1.1 Social Reform and Policy Research ...................................... 1
      The Policy Science Perspective ........................................... 3
      The Policy Analysis Perspective .......................................... 6
      Implications .................................................................. 11

  1.2 Program Evaluation as Policy Research ................................ 12
      Issues in the Theory and Practice of Program Evaluation .... 13
      The Program Evaluation Enterprise .................................... 16

  1.3 The Research Topic: Approach, Sources, Methods and Models 20
      The Research Problem .................................................... 20
      Theoretical Sources ....................................................... 23
      Methodology ................................................................ 26
      Models of Negotiated Knowledge Production and Use ....... 31

CHAPTER 2 CONTEMPORARY MODELS OF PROGRAM EVALUATION: PARADIGMS AND LIMITATIONS ......................................... 36

  2.1 Introduction ................................................................ 36

  2.2 The Agricultural-Botany Paradigm and Conventional Evaluations ..................................................... 37

  2.3 The Social Anthropology Paradigm and Illuminative Evaluations ..................................................... 41

  2.4 Limitations and Criticisms from Within Each Paradigm . 47
      The Interpretive Critique of the Positivist Model ................. 48
      The Positivist Critique of the Interpretive Model .............. 52

  2.5 Implications ................................................................ 55

CHAPTER 3 THE NEGOTIATED ORDER PARADIGM AND AN ANALYSIS OF PROGRAM EVALUATION CONTEXTS ................................................. 61

  3.1 Introduction ................................................................ 61

  3.2 Overview of Strauss on Negotiation ................................. 63
      Negotiation and Social Order ............................................. 63
      Background to the Negotiated Order Paradigm .................. 65
      Key Concepts in the Paradigm ........................................... 66
      Strauss' Use of the Paradigm in Case Study Analysis .......... 69
      The Strauss Paradigm: Summary and Implications .......... 71

  3.3 The Proposed Analytical Framework for Policy Research Studies ..................................................... 75
      Background to the Framework ............................................ 75
      Analyzing Structural Context ............................................ 79
      Analyzing Basic Areas of Negotiation ............................... 83
      Implications for Case Study Analysis and Model Development ..................................................... 99
CHAPTER 4  CHANGING STRUCTURAL CONTEXT AND PATTERNS OF COOPERATION AND ANTAGONISM .............................................. 101

4.1 The Structural Context: The Policy Shift from "Slum Clearance" to "Neighbourhood and Housing Rehabilitation"........... 102  
Policy Issues: Production vs. Distribution of Housing ............ 103  
The Evolution of Rehabilitation Policy ................................. 106  

4.2 Case: Stage II Evaluation of the Strathcona Rehabilitation Project ......................................................... 108  
Background to the Stage II Evaluation ................................. 111  
Phase 1: Pre-Contract Negotiations ..................................... 112  
Phase 2: The Evaluation Research ....................................... 115  
Phase 3: The Evaluation Report .......................................... 120  

4.3 The SRP Stage II Evaluation: Analysis of the Relations Between Structural Context, Negotiation Context, and Policy Consequences .................................................. 129  
Basic Areas of Negotiation .............................................. 131  
Summary and Implications for Knowledge Production and Use ......................................................... 138  

4.4 Case: Evaluation of the Neighbourhood Improvement Program Delivery System in British Columbia ....................... 142  
The Structural Context of the NIP/RRAP Program .................. 142  
NIP and RRAP: Citizen Evaluation ....................................... 144  
The British Columbia Evaluation: A Brief Chronology ............. 146  
The Evaluation Process: Methodological Conflict ................. 150  
An Ethnography of Conflict: The Details of Interaction .......... 158  
The Fieldwork Process: An Exercise in Research Team Conflict ......................................................... 159  
Organization of the Final Report and The Action Plan Meeting ......................................................... 162  

4.5 The NIP/RRAP Evaluation: A Study in Ambiguity and Antagonism ......................................................... 163  
Basic Areas of Negotiation .............................................. 164  
Summary and Implications for Knowledge Production and Use ......................................................... 169  

CHAPTER 5  PROGRAM SCALE, LEVELS OF INVOLVEMENT AND THE SUBSEQUENT USES OF EVALUATION STUDIES IN ONGOING PROGRAMS .................................................. 176  

5.1 Case: An Assessment of the Vancouver Jewish Family Service Agency ......................................................... 177  
The Structural Context: Public/Private Social Service, Community Organization, and Patterns of Charitable Contributions ......................................................... 178  
Preliminary Negotiations: The Outside Evaluators as Determined "Outsiders" ......................................................... 182  
The Research Process: Methodology, Short-Term Consequences and Problems of Distribution ......................... 183  

5.2 The JFSA Assessment: Levels of Involvement, Evaluation Use, and the Legacy of an Unchanged Structural Context ........... 187  
Basic Areas of Negotiation .............................................. 188  
Ongoing Problems and the Use/Misuse of the Report ............... 192  
Summary and Implications for Knowledge Production and Use ......................................................... 195  

5.3 Case: Legal Services Society Schools Program Evaluation Report 1980

The Structural Context: Stability and Shifts

The Schools Program Evaluation: A Brief Chronology

Defining the Process: The Decision, Type of Evaluation and Consultant's Role

The Internal Evaluation: Evaluation Needs and Methodology

Implications of the Internal Evaluation and Its Methodology

Subsequent Aspects of Program Change

Assembling the Evaluation Report

5.4 The LSS Schools Program Evaluation: The Locus of Evaluation Ownership and the Process of Organizational Change

Basic Areas of Negotiation

The Impact of Structural Context on Long-Term Change

Overview and Implications for Knowledge Production and Use

CHAPTER 6 MODELS OF NEGOTIATION KNOWLEDGE PRODUCTION AND USE: SOURCES AND DEVELOPMENT

6.1 The Theoretical Approach and Relevant Debate

The Analytical Strategy

The Choice of Analytical Framework and Relevant Debate

6.2 Overview of the Case Studies: Implications for Knowledge Production and Use

Studies in Housing Policy

The SRP Stage II Evaluation: Conclusions

The NIP/RRAP Evaluation: Conclusions

Implications for Knowledge Production and Use

Studies in Private Social Services and Legal Education

The JFSA Evaluation: Conclusions and Implications

LSS Schools Project Evaluation: Conclusions and Implications

Summary of Four Evaluation Case Studies

6.3 Preliminary Formulations: Models of Negotiated Knowledge

Background to the Models

The Experimental Model

The Managerial Model

The Collaborative Model

The Transformative Model

6.4 Possible Directions for Future Research

BIBLIOGRAPHY

APPENDICES
LIST OF FIGURES

FIGURE 1: Card with List of Six Eligibility Criteria .................. 155
FIGURE 2: Organization Chart of LSC 1978/1979 ....................... 201
FIGURE 3: Organization Chart of LSS 1980 .............................. 202
FIGURE 4: Summary of Four Evaluation Case Studies ................... 247
ACKNOWLEDGMENTS

I would like to thank the Chairperson of my academic committee, Dr. Elvi Whittaker, for her assistance as I struggled through several analytical approaches before constructing the particular documentary reality presented here. I have also benefitted from Dr. Robert Ratner's critical discussions of the internal logic of the work and clarification of the connections between each stage in the analysis. Dr. Roy Turner and Dr. Neil Guppy, from very different perspectives within sociology, led me to look more closely at the objections to the negotiated order paradigm, and marshal my arguments more systematically.

I want to congratulate my husband Seymour and daughter Rachel, who have endured and survived with few complaints and endless support. To Diana Raynor and Theresa MacFarlane, my typists, I owe much gratitude for their willingness to cope with the inevitable changes and details of each draft.
PREFACE

There are two characteristics of this research which contribute to its uniqueness as a study of program evaluation as a form of policy research. First, this work uses the methods and techniques of sociology and social anthropology to examine the activities of social scientists doing a type of field research. Thus, this description and analysis emphasizes the details of program evaluation work from the perspective of the evaluators as actors in various social settings rather than experts who do their work outside the flow of day-to-day social action. In addition, the case studies discussed represent a secondary analysis by me of program evaluation studies in which I was a major participant. That is, I am examining from a fresh viewpoint field data and research products that were generated by me, in cooperation with other researchers, for use in four different policy settings.

It is a bit difficult to specify all the methods and sources employed in producing this dissertation. It is essential that readers understand that during more than three-fourths of my work life I have been employed in non-academic settings doing applied and/or action research. These jobs have involved me in a wide variety of communities (geographic, ethnic, economic) and were generally connected with one or another type of "social reform" program.

The actual programs in which I have been employed ranged quite widely in scope and content. They include: assisting in an experimental program for high I.Q. pre-adolescent delinquents in Chicago; planning for new integrated school districts in Providence, Rhode Island; evaluating the first year's results of Operation HeadStart in Massachusetts; working as an urban planner and then coauthoring a manual on the use of survey research methods as a technique for improving the process of comprehensive urban planning (for the
Department of Housing and Urban Development in the United States); seven years continued involvement (as Vancouver Field Staff for the Company of Young Canadians) in the organization and development of a series of Vancouver self-help groups, e.g., Women's Health Collective, Mental Patients' Association, CCEC Credit Union and several geographically based community groups. In the most recent phase of my work life I have been self-employed as a planning consultant and the coauthor of several program evaluations, including the four which are used as case studies in this dissertation. One other unusual consulting contract involved me in the production of a market feasibility study for the Vancouver Native Village Development Society, a British Columbia Native Indian group interested in the development of an educational and entertainment complex in the Vancouver harbour area.

This employment has required that I work in a variety of occupational roles: (1) as an "organizer" of demands for social reform, (2) as a "designer" of social reform programs for specific social/political constituencies, (3) as an "implementer" of other people's reform programs, and (4) as the "producer of evaluations" of my own and other people's programs. In all four types of roles I was required to do evaluative work regarding the programs themselves and the personnel employed by them. It is important to point out that the approach I take when working "in the field" has some relationship to the methods I have used in interpreting my field experiences (and many other types of resources) for the analysis presented here. Without becoming too philosophical, let me say that in my work as researcher/organizer/evaluator I follow a kind of "naturalistic behaviourism". In this version of the research act one is involved in "the studied commitment to actively enter the worlds of native people and to render those worlds understandable from the standpoint of a theory that is grounded in the behaviours, language, definitions, attitudes, and feelings of those studied" (Denzin 1971). I very much appreciate Denzin's
expansion of the term "native person." He says, in a footnote, that "native person" is "introduced here as a generic term which covers all people studied by the sociologist. They may be experimental subjects, interviewees, friends, colleagues, and neighbours." In this sense, then, much of my work is firmly based in the concepts of everyday life, day-to-day activities, etc., as the source of theoretical insights and the initial formulation of evaluative constructs. But, along with Denzin, I have always been prepared to use (and often do employ) the full range of quantitative techniques in doing evaluative research. The particular combination of methods depends on their appropriateness to the requirements of a specific research problem.

Certainly, there are strong elements of that type of naturalistic behaviourism involved in my efforts to produce this dissertation. First of all, the research topic and methods reported here did not start with my receiving extensive training in the prevailing theoretical notions about how evaluation research is/should be accomplished. It was only after three years in the field that I began to read the abundant literature on program evaluation (as art and science). Up to then I had relied on several years of graduate training in sociological research and my work in urban planning and community organization as the basis for doing program evaluations.

During the initial years that I became a partner in Ptarmigan Planning Associates, self-employed as a program evaluator, I was sensitive to the risks and complexities of the work. As a result of previous work and training I was constantly aware of the variations in the actors, issues and settings around which evaluation research was being done. Each program presented a new experience in working out different arrangements regarding the planning, implementation and write up of the research and recommendations. Certain fundamental problems in accomplishing the required evaluative work were defined in alternative ways in each research setting. These included problems regarding
structural arrangement in which the work would be done; the conceptual perspective from which the evaluation process would be developed; the methodological strategies which would be employed; the format for document production, publication and distribution; and the extent to which the evaluator would be engaged in policy formulation and implementation. All this entailed a process of stating, discussing and trying to resolve a wide range of differences among the active participants (managers, staff, politicians, leaders of community groups, government representatives, concerned clients and the evaluator or evaluation team) in the evaluation activities. During the course of working on four program evaluations I came to see this process as a variant of what is traditionally defined as negotiation.

It should be understood that, in contrast to my emphasis on "negotiated" arrangements, several of the more popular texts and handbooks on program evaluation (Rossi and Williams 1972; Boruch and Riecken 1975; Guttentag and Struening 1975) tended to give a fairly rigid, experimentally based prescription regarding "good" evaluation research. Their models left out many of the essential actors in the evaluation process. The varied patterns of interaction among sponsors, researchers, managers, staff, participants (clients, patients), politicians and policymakers were often subsumed within a residual category called "research complications/barriers." These complexities of economic/social/political interaction in the field were often construed as obstacles to achieving one's main goal: a rigorous, objective evaluation of a specifically describable program in social reform. After brief discussions of strategies designed to overcome such complications, many of the texts gave their main attention to a detailed explication of scientific standards which must be met in order to produce an objective evaluation. Such books often concluded by outlining what one unfortunately had to settle for in place of rigorous, objective evaluation designs. These contrasting conceptualizations of the
experience and meaning of the program evaluation process (i.e., my own and the
version in the conventional text) were the source of my concern with documenting
actual variations in doing this particular form of social research.
Looking back, we can see that time and again political passion has been a driving spirit behind the call for rational analysis. On the one hand, rationalists advocated bringing to social affairs the neutrality and detachment that have served natural science and technology so well. And, on the other hand, the very call for value-free inquiry was politically motivated, generally by discontent. The savants of the Enlightenment were fighting traditions entrenched in royal and ecclesiastical power. The Progressives wanted social initiatives to which the legislators and mayors of the time were indifferent. The policy makers of the 1960's sought to use analytical methods to break up bureaucratic hardpan so that new programs could take root (Cronbach 1980:35).

CHAPTER 1: SOCIAL REFORM, POLICY RESEARCH AND THE PROGRAM EVALUATION PROCESS

1.1 Social Reform and Policy Research

Social reform programs have a long and respectable history in Western liberal democratic tradition. Ever since the beginning of the nineteenth century there have been significant programs to study and improve social conditions (e.g., sanitation, housing, health care) particularly among the urban poor. Many individuals and organizations were concerned with the systematic planning and implementation of the "reform" of the problems of urban society. One classic example was the work of Sydney and Beatrice Webb and other members of the Fabian Society in England. Their social and economic research was designed to provide a systematic knowledge base for improvements to the housing and welfare programs in Britain.

In recent decades, in both Europe and North America, the size and scope of the programs have increased quite dramatically. At the same time, the initiative for social reform has been co-opted into the political/economic
institutions of the State. In the U.S. new projects for social reform originally funded and promoted by the Ford Foundation and the President's Committee on Juvenile Delinquency (during 1960-1964), were ultimately incorporated into the American "War on Poverty" (Marris and Rein 1972). Similar innovative programs affecting housing, neighbourhoods, employment and health care were initiated in Canada with high hopes for the solution of basic urban problems (Dennis and Fish 1972; Draper 1971). When these reform programs did not fulfill their publicized expectations, all types of reform programs were subject to much closer scrutiny than in earlier decades. Decreases in economic growth rates, rapid inflation, and limited sources for increasing government revenues encouraged the policy makers' demands for systematic evaluation of reform programs funded by all levels of government. As a consequence, more "scientific" planning, monitoring and evaluation of these programs has become institutionalized as part of the overall policy intervention and program funding process.

The initiative for co-opting social reform programs into the structure of the State was part of an overall process of government expansion and required parallel development of an area of government activity called public policymaking. The notion of "public policy" has been in use for the last two centuries at least (Cronbach 1980:25), but it seems to have been reified as an entity, institution or object of scientific analysis in the last twenty-five years. Starting from an initial base in sociology and political science there have been two major perspectives employed in the effort to scrutinize the potential contribution of social science theory and knowledge to the formulation of public policy. The two perspectives may be labelled "policy science" and "policy analysis."
The Policy Science Perspective

The policy science perspective directs social science research toward rationalizing the process of making public policy (Lazarsfeld 1975). Etzioni has referred to policy science as an "in-between discipline" (1971:11), designed to bridge the gap between abstract theory and policy recommendations. Thus, policy science furnishes "practical knowledge that has no analytic foundation but is useful for moderating or complementing analytic knowledge" (Etzioni 1978:4). The emphasis in policy science is to "prepare alternative rations and to pry policymakers loose from their antiquated assumptions" (1978:2). From these comments it becomes evident that one of the main assumptions of policy science is a separation of the value component of social research from the fact finding activities of the social researcher. From the policy science perspective this often takes the form of a sharp distinction between ends and means, between the process of setting policy goals and the research activities entailed in the planning, implementation and evaluation of those policies as they are interpreted in specific publicly funded programs. Policy science is critical of the limited research capabilities of most policymakers and their staffs, arguing that the usual task forces or commissions "frequently lack the facilities, experience, and continuity--in short, the structural prerequisites for conducting high quality research" (1978:3). What seems necessary are more facilities on the domestic front (like RAND Corporation working for the armed services) in which policy research is institutionalized.

As a consequence of these concerns, major attention has been given to policy research, which is broadly defined as "decision oriented research" (Coleman 1972:2). "More particularly policy research is "concerned with mapping alternative approaches and with specifying potential differences in the intention, effect, and cost of various programs" (Etzioni 1978:1). The work of policy research is seen as a distinctive enterprise with purposes that
differ from both pure and applied social science research. In principle policy research is defined as being more critical than applied research (Etzioni 1971:9), while being less formalized, less analytic than pure research (Etzioni 1978:4).

Following the approach taken by Coleman and Etzioni, an effort has been made to outline specific characteristics of policy research. Finsterbusch and Hamilton (1978:88) offer the following:

1. the research problem originates in the policy arena;
2. the audience for the research results is a group of political actors and interested parties, usually with conflicting interests;
3. the research must examine the consequences of a policy for all interested parties;
4. the research is constrained by the timing of decisions;
5. the policy researcher must invest a significant amount of time and energy in communicating with policymakers and interested parties.

Examples of policy studies include social impact assessments, studies for environmental impact statements, reports to various policy oriented commissions and program evaluations.

In this context policy research is seen as offering a set of rational procedures and results which can be used in policy formulation as well as the planning, implementation and evaluation of public programs. "The deviation from functional rationality in social science research begins with the failure to clearly define the overall research objective and the system of instrumental subgoals" (Finsterbusch and Hamilton 1978:89). Although goal setting is seen as a continuous, iterative process, there is great stress on goal definition and the legal/political context within which the research results will be received and processed.

However, not all those concerned with policy science, and policy research in particular, take such a strong positivist line. In their analysis of two studies of Community Mental Health Centers, Siegel and Doty (1978:144) suggest
that policy research studies can vary considerably in their emphasis on goals and purposes. Their two case studies differed significantly, both in their research orientation and structural relationship to the program being studied. The "management analysis", conducted internally by a knowledge unit which was an integral part of a policymaking body, took purpose and implementation strategy largely as a "given". The "advocacy research", produced by independent outsiders claiming to represent the interests of the community not the program, questioned both program purposes and implementation strategy (1978:140). In analyzing these two pieces of policy research Siegel and Doty bring the policy science approach closer to that of policy analysis. That is, their analysis focuses on the balance of "intrinsic" (paradigm modification) to "extrinsic" (paradigm shift) criticism of the policy and practices of the Community Mental Health Centers presented in each of the two case studies. However, the authors continue to be interested in the goal analysis developed in each study, criticizing the procedures and limitations of the methodologies employed in both.

Etzioni stresses the critical orientation of policy research, but sees no intrinsic difficulties in using the time, training, books and computers of social research to satisfy the information needs of all interested parties engaged in the policy process. In commenting on the most notorious phrase written into the 1964 anti-poverty legislation, "maximum feasible participation of the poor", he is extremely critical of the lack of preliminary investigation of what such participation entailed, how it was to be effected, and what the consequences were likely to be. Thus, he urges that:

If well planned policy research had been done before the enabling legislation was passed, the best methods of studying the problems and alternative solutions would have been delineated. The researchers, not necessarily accepting the explicit and implicit goals of the sponsors of the legislation, would have explored the underlying assumptions, examined the possible consequences of different alternatives, and evaluated the effects of small-scale demonstration projects (1978:2).

The approach of policy science, as a discipline that falls between the analytic sciences and policy recommendations, "the way medicine comes between physiology and pharmacy" (1978:4), would produce research results and "diagnoses" that could be effectively communicated to policymakers and other interest groups. Therefore, "knowledge makers who care should make their work relevant to the citizens at large and the policy makers of the institutions and movements which respond to their true allies" (1978:5). In this knowledge production framework there is a tacit assumption that policy research is essentially critical but simultaneously "value-free", i.e., that it can serve all potential users equally. The facts established by policy research are separate from the value component embedded in the goal setting activities of the policy process.

The Policy Analysis Perspective

Generally the "policy analysis" approach to the possible relationship between social science and public policy has been concerned with issues and problems closer to the center of political and social theory. This perspective can be seen in the work of Peter Marris and Martin Rein (1974) which documented both policy and program dilemmas which evolved in the community action and poverty programs of the 1960's. In his essays on social science and public policy, Rein (1976) suggests that the study of social policy is "basically concerned with the range of human needs and the social institutions created to meet them" (1976:20). But he emphasizes that we have no adequate definition of "need", and there is confusion about the distinction between "need", "preference", and "social problem". Furthermore, we find that the institutional arrangements designed to meet these needs appear to vary significantly, according to the historical, cultural and political setting in which they have been developed. Thus, the policy "truths" of one place or one decade may be challenged in
another. In Canada, for example, policymakers in Parliament and at Central Mortgage and Housing Corporation were committed to the principle that the "urban blight" of the inner cities had to be removed through extensive demolition of substandard housing to meet the need for adequate, safe and sanitary housing for all Canadians. The social, cultural and economic fallout from this policy was disproportionately heavy for low income and elderly urban residents. These target groups were subsequently housed in dense, high-rise projects which had their own negative impacts, i.e. alienation, increased levels of crime, violence and mental illness (Fried 1967; Dennis and Fish 1972; Clairmont and Magill 1974). Ultimately, in the 1970's there was another policy shift and programs for neighbourhood improvement and housing rehabilitation were funded as the appropriate institutional arrangement for assisting the poor, elderly and others on fixed incomes to meet their basic need for safe, clean and secure shelter. This example illustrates Rein's view that "ideas are embedded in practice; the relationship between means and ends of social policy is not settled, nor is it ever likely to be finally resolved" (1976:21). Evolving social objectives are central to the development of social policy; therefore policy analysis necessarily forces us to carefully examine the multiple, ambiguous and conflicting objectives of any specific policy.

Similarly Rein attempts to deal directly with the fact-value dilemma and offers a view of policy analysis as necessarily involving the interpretation of beliefs and research paradigms. He is critical of the positivist separation of facts, theories and values and delineates his argument as to why policy conflict cannot be resolved through reliance on policy-relevant social science. Since this argument assumes that facts and values are always integrated and cannot be isolated from one another (1976:79), he insists that we must become aware of the ideological perspectives that are typically encountered in the analysis of public policy. These "value screens" serve to "screen in and out certain evidence" and provide a framework for interpretation (1976:80). Rein
suggests three major value screens which can orient the policy analyst to different causes and means of resolving contemporary problems. These include: the policy analyst from the LIBERAL political tradition who will find that the institutions and organizations in society must be reformed; the CONSERVATIVE who calls attention to the failure of people to cope and finds people accountable for their own failures and achievements; and the RADICAL who sees fundamental change inhibited by the distribution of established power, so that important reform is possible only through revolution (1976:84). Not all ideologies emphasize one of these themes specifically and most policy analysts recognize some merit in each view, so that a synthesis can be organized. Nevertheless, Rein argues, one is forced however reluctantly to make choices, since each offers a competing framework of thought:

All these ideologies direct our attention to specific facts, and they also help to organize them. In this sense all research is biased: it works within an interpretive framework enabling the analyst to make policy-relevant sense of the facts that he observes (ibid:84).

As indicated previously, the policy science approach seeks to rationalize the policy process through the utilization of relevant social science techniques, employing concepts that reduce the need for transition and translation, and knowledge that is, in general; "not far removed from the observed world and that experienced by the policymaker" (Etzioni 1978:4). In contrast, Rein sees major obstacles which appear to frustrate the potential contribution of social science theory and knowledge to policy formulation. There is, first of all, what he sees as an inherent conflict in the political decision-making process, i.e., "different people want different things and make use of the political process to satisfy their respective interests or 'disinterested' aspirations" (Rein 1976:98). Second, even if there is no immediate clash of interests, implementation of research findings may still be a problem because of the complexity of the situation for which the policy is proposed (1976:101). For
example, Canada Manpower and Immigration funds training programs to bring marginal workers (young, poorly educated, ex-alcoholics, natives, women) into the labour force. Social researchers (Enoch and Levitan 1981) evaluate a woodwork training program as being relatively effective in preparing selected individuals for employment in kitchen-cabinet-making businesses and for entry into pre-apprenticeship programs run by the B.C. Carpenters' Union. However, the unemployment rate among fully qualified journeyman carpenters is close to 20%. How relevant is that program evaluation to questions of manpower policy in a depressed employment market which must be faced by marginal and fully qualified workers alike? Finally, Rein argues, in a pluralistic system with competing interpretive frameworks, social science findings are seldom so conclusive that they permit a firm choice—negative findings corroborate already known criticisms, while positive findings can be shown to be too "costly", or the product of careful selection of candidates likely to "succeed" (1976:106).

Yet, despite these obstacles, Rein takes an optimistic view, suggesting several research strategies through which empirical findings may contribute to the development of policy. One strategy, which is labelled a CONSENSUAL APPROACH, can be seen as the position most commonly taken by social researchers who follow the policy science approach to policy research advocated by Etzioni. In this context government contracts for the information it wants, e.g. the "management review" of Community Mental Health Centers done by the U.S. Government Accounting Office; or intermediaries bring together knowledge producers and knowledge users, e.g., Westinghouse Corporation brought together a group of consultants in educational testing, psychology and sociology to provide evaluation studies of Project Headstart for the Department of Health, Education & Welfare. Another strategy, which Rein has referred to as a CONTENTIOUS APPROACH, may be taken in settings where research acts as a MORAL WITNESS for the failure of society to honour its commitments, e.g., the
Stage II Evaluation of the Strathcona Rehabilitation Project fully documents serious procedural obstacles to citizen participation in the housing rehabilitation component and overspending on an excessively large capacity sewer and water system. Similarly the contentious approach is reflected when the researcher acts as a SOCIAL CRITIC under circumstances where government has established no commitments to be honoured, e.g. the various research studies done by women's groups and addiction foundations regarding women's use of minor tranquillizers and alcohol, prepared as part of a five year effort to obtain funding for necessary education and treatment programs. The third research strategy in which empirical findings contribute to the development of policy Rein defines as a PARADIGM-CHALLENGING APPROACH. This strategy requires the researcher to act independently of the established framework, to expose its fundamental weaknesses and propose alternative principles of intervention, while assuming that political support for the assault will follow at a later time (1976:125). There are limited examples of the latter approach; however, aspects of Kleiber and Light's (1978) descriptive and evaluative study of the Vancouver Women's Health Collective exposes certain fundamental weaknesses in the medical model of health care delivery to women and proposes extension of the self-help model to major areas of health care for women.

It is important to note that "while research is typically undertaken within one or another of these frames of reference, it is not uncommon for a single study to shift from one strategic level to another to accommodate the varied motives of the researcher" (1976:125). This shifting between levels becomes evident in any detailed analysis of empirical research which hopes to be useful in the development of policy. For example, the evaluation of Community Mental Health Centers in the U.S. done by a task force of the Center for the Study of Responsive Law ("Nader's Raiders") employed a rather traditional consensual approach to the design of the evaluative research. The authors kept to official documents in their description and analysis of program goals, did little
or no analysis of actual program delivery and never questioned tacit definitions of "community". However, in their final appraisal of the program they produced paradigm-challenging recommendations regarding the directions for action which should be followed in new mental health service delivery (Siegel and Doty 1978: 146).

With these provisos regarding obstacles likely to frustrate potential research use, and the problems of shifting strategy levels, Rein outlines three main areas of study which seem to be relevant to developing a viable social policy paradigm. These areas include:

1. documenting the disparities between existing ideals and established realities by empirical study and, as a member of society, using one's intuitive sense to assist in identifying which disparities may be most acceptable politically (1976:127);

2. studying theories of policy intervention in an action framework (ideologically and empirically linked to real social processes) designed to assist in reducing the disparity between ideals and reality (1976:130); and

3. evaluating specific programs which translate a broadly defined theory of intervention into concrete actions (1976:133).

Implications

Although the policy science and policy analysis approaches differ in their underlying epistemologies and theoretical statements about the potential uses of policy research, there are several areas in which they tend to coincide. First, both recognize the direct and immediate linkages that must be forged between the policy researcher and the various actors in public settings to which she must relate. Second, an acceptance of controversy is explicit in both approaches, although there is disagreement about the possibilities for resolving such conflict through research and analysis. Third, although the leading figures in the two camps stress the critical nature of policy research, there is an overall sense of optimism conveyed by both regarding the ultimate
usefulness of social science research in the policy process. And, finally, both approaches define the evaluation of government-funded programs concerned with social reform as a key area of policy research. The main theoretical and empirical work of this dissertation focuses on the program evaluation process as a form of policy research. However, at various points where it appears necessary to the analysis, this study will touch on questions of policy ideals: vs. the realities of legislation, implementation and funding, and explore the theories of policy intervention which underlie the goals and organizational arrangements of a particular program.

1.2 Program Evaluation as Policy Research

Program evaluation as a form of policy research can refer to a wide variety of descriptive and analytical studies of program process (formative evaluations) or program impact (summative evaluations), or some type of comprehensive program assessment which includes both modes of evaluation study. Cronbach offers the following definitions of "evaluation" and "program" which are sufficiently general as to allow a range of specific studies to be covered:

By the term EVALUATION, we mean systematic examination of events occurring in and consequent on a contemporary program--an examination conducted to assist in improving this program and other programs having the same general purpose. By the term PROGRAM, we mean a standing arrangement that provides for a social service (1980:14).

Following Cronbach's position (1980:17), I do not stress quantitative-statistical methods, or "goal attainment", or the intent to judge a program as good or bad in any definition of what constitutes a program evaluation study. Further, program evaluations may sometimes be concerned with an established program that is ongoing, sometimes with a program that had a limited time frame and has already been terminated, and in other instances with a plan that could
be established if found suitable. The main elements that program evaluation studies appear to have in common are that they are titled "program evaluation" and are generally designed to be used by a client or clients (other than the program evaluator herself) who have some vested interest in the program in question.

Obviously, I want to emphasize the diversity of policy research work that goes under the label program evaluation. But, what is just as important, I want to present an analysis of program evaluation as a process that involves the notion of "judging merit" (Weiss 1972:1), where merit itself is defined in diverse ways depending on the value position of the judge and the procedures selected for judgement.

**Issues in the Theory and Practice of Program Evaluation**

The recent literature on the use of social research in public policymaking (Rein 1976; Weiss 1977; Etzioni 1978) and program evaluation (Guttentag and Struening 1975; Franklin and Thrasher 1976; Cronbach 1980) suggests that thoroughgoing evaluation of an existing program or early planning for the evaluation of a newly developed program involves a series of theoretical and practical decisions. Such decisions require information about the specific program (this notion of "program" includes projects, agencies or sub-units of an overall nation-wide program) and the social/political context in which the program is operated. Any effort to examine empirically the specific activities entailed in constructing a program evaluation and implementing its results and recommendations requires some analysis of the general issues around which such decision-making tends to take place. These questions are dealt with in some detail in Chapter 3 (section 3.3), but I want to outline them here briefly since they become important in defining the topic of this dissertation.
The first issue which may become the focus of discussion in the program evaluation process is the delineation of all the major actors whose interests will be represented throughout the evaluation study. The formality (in contractual terms) or informality (as observers and/or critics) of that representation, and the scope and intensity of control over subsequent phases of the evaluation process may become controversial questions right at the start, and grow or fade as the work continues. A second, closely linked issue, is the question of the organizational placement of the program evaluation research. Decisions must be made by key actors whether the work will be done within the program setting (by "in-house" researchers), will be directed by program management using outside consultants, or be commissioned research done by a firm of program evaluators selected through a public process (requests for proposals, interviews, hiring committee, etc.). In large scale and/or federally funded programs the public process may be required, but the specific details in each case may vary significantly depending upon pre-existing implicit agreements between program funders, managers, and lists of preferred consultants.

The third issue, the general research strategy to be followed, and the fourth, the specific model and methodology to be used also have obvious connections. In analyzing questions of research strategy, we must understand that frequently this lies so far below the surface of discussion that it never emerges as an issue. However, there are settings in which an awareness of the "value screens" of the major actors, and overt consideration of the consequences of selecting the consensual, contentious or paradigm-challenging approach become the focus of serious analysis within the evaluation team and between that team and the funders of the evaluation study or program managers (Kleiber and Light 1978). More often, differences emerge around questions concerning the research model and/or specific methodology to be employed in the evaluation study. These frequently focus on the methods used in the research design, the.
positivists expecting formal experimentation and statistical measures, the ethnographers emphasizing the necessity for observational techniques, and those trained in the tradition of social analysis (Rowbottom 1977) seeking some balance of the two approaches likely to increase the use of the research findings.

The fifth issue around which there is often extensive discussion and negotiation in policy research studies generally (and program evaluations in particular), is the construction, content and timing of the research report. Reaching agreement on the acceptability of the final draft of a program evaluation may be simple and straightforward in some cases; whereas in others acceptance (and payment) may be withheld as a result of previously unresolved conflicts over models and methods, or over the content and interpretation of the research findings. Often the term "counterintuitive" is used in this context and usually it is a gloss for a situation where the expectations of the policymakers/program funders/managers have not been reinforced by the findings of the evaluation report. In these situations it becomes very difficult to find some way to discuss the sixth, and final issue, planning for research utilization. In most settings policy researchers make some systematic effort to assure that a "positive reception" is given to their research results and policy recommendations. However, where program evaluators have specifically chosen a contentious or paradigm-challenging approach their plans for research utilization may differ. In such situations the evaluation document is designed as a tool to provoke controversy and/or educate a constituency regarding its rights and needs. This issue of planning for research utilization may become the first phase of another cycle of program planning and change in those settings where the evaluation component is part of an overall management cycle (Crowfoot 1977).
These six topics (the major actors, organizational placement, research strategy, models and methods, research report and research utilization) are labelled in slightly different ways in the various major texts on program evaluation (Caro 1971; Carter and Wharf 1973; Cronbach 1980; Franklin and Thrasher 1976; Rossi, Freeman and Wright 1979). Each is given varying degrees of emphasis depending upon the authors' values, ideological position and research paradigm. However, as indicated in the preceding discussion, these research questions cannot stand on their own in any sociological analysis of the factors affecting the design, production and potential use of program evaluations as contributions to the policy process. It is necessary to examine the larger organizational arrangements in which policy research is conducted and to develop systematic documentation of how these research issues are linked to one another in specific program evaluation settings. An extensive examination of the social, political, economic and organizational aspects of the policy intervention process is presented in Chapter 3 (section 3.3). The only point I want to make here is that no model of policy research can be complete without systematic consideration of these aspects of the policy process being included as part of the strategy of program description and assessment.

The Program Evaluation Enterprise

Before proceeding with a discussion of my research topic, I want to present briefly some of the characteristics of program evaluation as a knowledge production industry. This background material is intended as one way to outline for readers the extent to which the work of program evaluation research is empirically linked to aspects of the larger political economy, both the public and the private sectors.
During the past decade managers of programs in the areas of health, education, social housing and welfare services were forcefully notified that their programs had to become more efficient as well as becoming more politically accountable. They had to look to a new kind of specialist, the program evaluator, to help solve the problems of demonstrating program efficiency and documenting program impact. From the perspective of these program managers there were many reasons for permitting and occasionally even encouraging an evaluation of their programs. From the vantage point of politicians and other policymakers there was a different, but often overlapping set of reasons for instituting a process of evaluation of all programs coming up for new and/or continued funding. The resulting evaluations were usually initiated for many different and, at times, even conflicting purposes; for example; choosing the best among several alternative forms of program development, re-adjusting the organization of a program already in operation, maintaining quality control over program services, forcing subordinates to comply with policies from above, documenting that one's agency (or agency supervised by one's policy committee) deserves its present budget or a budget increase, creating support for a favourite new project, casting suspicion on a policy supported by political opponents or a competing program, and so on.

From the perspective of the program evaluator the range of purposes was always greater than any single study could fulfill, even when managers and policymakers were direct and honest about their evaluation agendas. The social, political and economic contexts of this process varied so extensively that no individual was fully equipped to cope with all contingencies. New policies and new programs at all three levels of government presented a complex and growing market for the skills of program evaluators from all the social science disciplines. Furthermore, as the funding sources for social service, housing and education programs were reduced, as a result of inflation and
ideological shifts, such programs increasingly found it necessary to justify their impact and produce good "bottom line" results.

Program evaluation has become an enterprise that requires a team of researchers and, in the case of large scale government funded programs, the evaluation team requires substantial organizational backing from a firm or research institute. There is little "hard data" on program evaluation firms in Canada, and the main bases for this part of the discussion derive from sources relevant to the United States (Abt 1976; Cronbach 1980). In neither case, Canada nor the United States, do we have any systematic information on the number of firms nor a comprehensive assessment of the quality of the knowledge product these firms are paid to produce. In general, we have a sense that the demand for evaluations is tied directly to levels of government spending during each fiscal year, and the expansion/contraction cycle for individual firms can be very rapid (Cronbach 1980). There are a very small number of vertically integrated firms that obtain the large budget contracts (e.g., Peat, Marwick and Partners in Canada, Abt Associates in the United States), and a rapidly fluctuating number of small research groups and individuals who operate on a "piece-work" basis. The few large firms have full-time staffs of social scientists, networks of field observers and interviewers, public relations personnel and collections of specialized consultants on call (e.g., petroleum engineers, urban archaeologists). Several of the smaller firms and individuals subsist through subcontracting to larger firms, and/or having other sources of stable income. This is particularly the case for sociologists/anthropologists who have their economic base in university or college teaching and for planners/ecologists who are employed by architectural or engineering firms.

One additional aspect of the economic structure should be mentioned here, i.e., the increasing scale of evaluation studies. It has been estimated that
at the beginning of the 1970's, in the United States, about 300 new studies were begun each year, with direct federal support and with an average budget of about $100,000 (Eaton 1980). At the close of the 1970's requirements for large-scale evaluations had increased to a point where it was not uncommon for a single major study to have expenditures over one million dollars a year (Cronbach 1980:39). Both Cronbach (1980) and Eaton (1980) use the term "mushroomed" when referring to the unplanned growth of evaluation activities in the last ten years. It has been said that there is a "boom town" excitement in the evaluation community. But it appears that federal expenditures in the United States for evaluation studies have been nearly level in the last few years. A survey of policy research undertaken by federal government departments and agencies during 1975-1977 indicated that on an overall base of one billion or just slightly more, program evaluation represented between 5.2% and 5.6% of budget obligations in each of those three fiscal years (Cronbach 1980:42). Although we lack similar survey data for comparable federal government departments here in Canada, it appears safe to assume that the economic picture for program evaluation is rather similar. For example, in the last few years there has been a marked reduction in the amount available for social housing programs at CMHC and fewer dollars for innovative programs at Health and Welfare. The one area that has seen continued evaluation and social impact studies is primary resource development in the western provinces and the Territories. Finally, we should note that smaller United States and Canadian firms continue to assist local agencies and municipal governments to monitor and assess their programs. At the same time, the larger firms have begun to play an important role in the evaluation of programs funded by the federal level of government but developed for and implemented in various parts of the Caribbean, Africa and Southeast Asia (Weiss 1977).
1.3 The Research Topic: Approach, Sources, Methods and Models

Since every program evaluation is responsive to different aspects of the six theoretical/practical questions and yields a different set of inputs into the policy process, there is a danger that social scientists concerned with making generalizations about program evaluation as a type of policy research will stay with the dominant policy science paradigm. From this perspective one must take the experimental/positivist model as a research standard, construct a role for the evaluator as a relatively passive technician, and define many of the contradictions of the evaluation process as the products of inadequate communication, faulty methodology and sloppy statistical design. On the other hand, there is a second perspective available which includes elements of the policy science paradigm, but combines it with the philosophical and political insights of policy analysis. In this alternative the evaluator is seen as an active policy analyst and interpreter of reality, who has a number of research models and methods in her repertoire, and attempts to use a critical but essentially optimistic approach to the problems entailed in producing a program evaluation. Generally, the work of this dissertation takes its starting point in the second approach to policy research and program evaluation.

The Research Problem

The general focus of this research involves looking at the uses of social research in public policymaking, and, in particular, examining the construction and use of program evaluation studies in a series of social reform programs. In the course of setting up the research design for this dissertation I have made certain conceptual decisions which entailed the use of multiple, theoretically compatible paradigms: the sociology of
knowledge, social organization theory and a recent version of the negotiated order paradigm (Burrell and Morgan 1979).

I consider program evaluation studies and other forms of policy research as socially produced documents which are open to analysis through the use of a sociology of knowledge approach (Smith 1974b). These socially constructed documentary realities are "products" which are purchased by a definable organization whose management wishes or is required to fund such a program evaluation study. These specific organizational "consumers" of the research differ in their basic structures, decision-making orientations, management ideologies and several other factors to which organization theory is directly relevant. However, the consumers of the study may include other major actors outside the funding organization; thus, community residents, program clients, the media and other interested groups become members of a larger "market" for program evaluation studies. In order to study this larger structural context I have used a third paradigm to attempt to link up the production and consumption aspects of the research market. It is the negotiated order paradigm (Strauss 1978) that I have employed as a theoretical guide for integrating issues in the construction of the policy research document (the program evaluation study) with the problems of analyzing the organizational structure and operations of the consumers of such research.

In this effort to analyze the process, products, users and uses of policy research, I found that there were critical limitations in the research models that tend to predominate in discussions of program evaluation work. Therefore, I came to the decision that the specific purpose of this dissertation would be to develop a wider selection of evaluation research models which would specifically take into account construction of the research "product" and characteristics of the larger structural context in which such products
are designed to be used. My theoretical strategy involves an effort to relate the more stable, structural characteristics of the social, political, economic, and organizational contexts of reform programs to the series of six theoretical/practical issues (outlined in 1.2) in the program evaluation process. The expanded set of policy research models emerge from both theoretical and case study sources.

The first step in this process of formulating new models involves a critical review of the two major models of program evaluation research described in the contemporary literature on policy research in education. The critical review examines each model in terms of its main assumptions and directives regarding resolution of the six theoretical/practical issues. Second, taking a symbolic interactionist perspective on the study of social organization and social structure (the negotiated order paradigm), I construct a framework for the analysis of program evaluation studies which specifies components of the structural context of policy research and redefines the six theoretical/practical issues as "basic areas of negotiation". Then using this framework, I describe and analyze the program evaluation process in two pairs of evaluation research studies. The case studies provide an opportunity to explore empirically the differing patterns of knowledge production and use which emerge when varying decisions are negotiated involving the six policy research issues.

In the final chapter I outline and discuss several models of negotiated knowledge production and use suggested in the course of the case study analysis. These models are not designed to be either exclusive or exhaustive. Rather, they are presented in order to indicate (a) the range of negotiated options which may surround each of the major theoretical/practical research issues, and (b) the degree to which structural and negotiation contexts may
shape the policy research process. In addition, these models suggest possible directions for "sampling the population" of program evaluation studies to develop additional models of negotiated knowledge production and use in the policy research process.

Theoretical Sources

The most fully developed discussion of program evaluation as a form of policy research has emerged in the evaluation of schools and education programs generally. Therefore the most extensive, immediate source for an analysis of evaluation research is located in the literature on contemporary models of program evaluation developed and used in policy research in education. The discussion presented in Chapter 2 is intended as a critical overview of the principal models of research that educational program evaluators claim are most frequently employed in doing their type of policy research. Although these theoretical examples (or "ideal types") have been taken from educational research, the descriptive and critical remarks are directly applicable to examples of policy research in the areas of housing, health care, welfare services, or employment training. Substantive details of the research models may be shifted without finding significant changes being made in the underlying paradigms incorporated into the two models which are presented for critical analysis.

Description of the agricultural-botany paradigm and conventional program evaluation studies relies on the work of Rossi and Williams (1972), Boruch and Riecken (1975) and several articles in the Handbook of Evaluation Research edited by Guttentag and Struening (1975). The basic framework for this model of program evaluation derives from a positivist/experimentalist paradigm of social science. The alternative view of evaluation research which has
emerged in the last decade challenges many of the assumptions of the positivist model. This second view derives from an anthropological or interpretive paradigm of social science research and has been termed "illuminative" evaluation. The discussion of the second model relies on work by Parlett and Hamilton (1976), R.E. Stake (1974), and commentary by the Stanford Evaluation Consortium (1976).

In Chapter 3 my focus shifts to the problem of constructing a theoretical framework for the analysis of program evaluation research which extends beyond the limits of the models presented in Chapter 2. After having reviewed two models of how program evaluation should be accomplished, I direct my attention to a theoretical perspective which encourages analysis of program evaluation research as it is accomplished in specific organizational contexts. The discussion presented in Chapter 3 relies on the theoretical perspective of symbolic interaction, and more specifically the negotiated order paradigm developed by Anselm Strauss. The central idea in this theory of social organization is a concept of social structure as being "in process", and the belief that "social orders are always in some sense negotiated orders" (1978:xii). Strauss describes the social order in terms of the process of ongoing negotiations which is carried on by many types of participants, each with distinctive world views. Negotiations are conducted simultaneously at several levels in the social order, and each level can offer resources to and impose constraints on others.

In developing the negotiated order paradigm Strauss defines two key terms: "structural context" and "negotiation context" (1978:99). The structural context refers to salient structural properties that bear on negotiation, the more stable arrangements within which negotiations take place. The concept of negotiation context is analogous to the use of the
term "awareness context" in his earlier work (Glaser and Strauss 1964). He specifies that there are different kinds of negotiation contexts which pertain to interaction among negotiating parties. These are related to permutations of several properties identifiable in any negotiation context: the negotiators, the nature and timing of the negotiations, the balance of power, the nature of the respective stakes, the visibility of the transactions, the number and complexity of the issues, the clarity of legitimacy boundaries, and options to avoiding or discontinuing negotiation (1978:100).

He remarks that:

A steady focus on structural and negotiation contexts and on their respective properties increases the likelihood that the analysis of specific courses of negotiation will be carefully located "within" the larger social structure. (No reification is intended regarding social structure) (1978:101).

This paradigm is directly applicable to an analysis of the evaluation research process and uses of the documentary product in the planning for program stability and/or change. Its main purpose is to provide a theoretical guide for integrating elements of the two other paradigms involved in this analysis of the construction and use of policy research: the sociology of knowledge approach used in examining the process and product of research (the evaluation document) and the organization theory employed in the effort to specify the organizational consumers of such research (program funders, managers and staff, policymakers and other interested parties). Used as a means to make theoretically grounded observations, the negotiated order paradigm helps to characterize the negotiation context of a specific program evaluation in relation to the overall structural context of the policy intervention approach which served to guide program development and implementation. Second, it offers a framework for assessing what has been omitted or glossed over in existing models of program evaluation research which tend to focus
on selected properties of the negotiation context and screen out elements in the structural context which may offer resources/constraints on every aspect of the evaluation process. Finally, it permits us to conceptualize each of the six theoretical/practical issues of policy research as a negotiation context for interaction among several negotiating parties. This suggests a wide variety of research questions which generally go unasked in the two main models of program evaluation research. For example:

Under what circumstances might all the major actors in the program evaluation process become actively involved in the field research and planning for research utilization?

Does the research methodology become a source of conflict among the parties only at the design stage or can it re-emerge as an issue subsequently, especially where research results are defined as being "counterintuitive" by policymakers?

How might differences in overall research strategy affect negotiations among various evaluation teams in large scale evaluation studies?

Do key actors in the structural context of the program become involved in negotiations regarding research strategy or subsequent utilization of research findings?

These are examples of the kinds of detailed research questions which provide clues to contextual differences in the production and use of policy-relevant research knowledge. The negotiated order paradigm offers a theoretical framework which encourages close empirical scrutiny of both the structural and negotiation contexts of each program evaluation, and careful documentation of the distinctive world views held by key participants in these negotiations.

Methodology: Case Studies Based on Documents, Interviews and Field Data

In Chapters 4 and 5 I describe and analyze two sets of program evaluation case studies in terms of their structural contexts, the basic areas of negotiation and differing patterns of knowledge production and use which emerged
in each specific evaluation context. The first two sets of case studies were evaluations commissioned by various levels of government (and in one case a community group also) to assess the delivery of service and overall program impact during a shift in Canadian housing policy. The original intervention theory had focused on the goal of slum clearance, while the new policy was oriented to a goal of upgrading the conditions in existing housing stock and rehabilitating the physical facilities and amenities of local neighbourhoods. The initial case study presented in Chapter 4 concerns the evaluation of the Strathcona Rehabilitation Project, a housing and neighbourhood rehabilitation pilot project located in the Strathcona area, a neighbourhood adjacent to the central business district in Vancouver. The SRP, as one of a series of pilot studies, was designed to serve as a model for subsequent national legislation regarding housing conservation and neighbourhood rehabilitation in cities and towns across Canada. The second case study discussed in Chapter 4 concerns an evaluation of the new national legislation as it was implemented in British Columbia, i.e., evaluation of the Neighbourhood Improvement Program (NIP) and Residential Rehabilitation Assistance Program (RRAP) as they were delivered in selected urban neighbourhoods and rural towns in B.C. Although the two case studies involve the same substantive policy area (housing), overlapping time frames, and a similar research team, there are significant differences in the structural and negotiation contexts of the two evaluations which emerge in their descriptions and analyses in Chapter 4. These variations offer the initial indicators of the differing patterns of knowledge production and use in policy research studies.

In Chapter 5 I present two additional program evaluation case studies in order to extend my description and analysis of the program evaluation
process. The first study described and analyzed in that chapter focuses on an assessment of the Vancouver Jewish Family Service Agency, a private social service agency which is supported by funding from the United Way and the Jewish Community Fund and Council. The second study concerns the evaluation of the Legal Services Society Schools Program, an ongoing program for the delivery of legal education services and materials in the B.C. school system, supported by funding from the Provincial Attorney General's Ministry but accountable also to the Ministry of Education (which provides indirect funding through release-time for teachers, some transportation assistance, etc.). These two evaluation case studies differ from each other, as well as differing from the cases presented in Chapter 4, in terms of several organizational/structural factors, e.g., substantive policy area, scale of operation, their ongoing nature and length of time in existence. However, despite overt differences in the two programs, the case study analysis suggests important similarities in the structural contexts of each program which may have contributed to the pattern of knowledge production and use of the evaluation research materials.

Since I had an ongoing interest in questions regarding the construction of research documents and the uses of these research findings by program funders, managers and policymakers, I began collecting field data on each of the program evaluation studies (as well as other studies done by members of the Ptarmigan Planning work group) at the time of the initial contact with members of the organization making the research request. The main methodological approaches which guided the data collection were designed to serve both the specific purpose of producing an evaluation document and the more general purpose of describing and analyzing the process of document production and use in policy settings. Each of the contracts or letters of agreement
with funding organizations gave permission for the research data and the resulting document to be used for academic purposes. In order to reach conclusions which had a relatively strong level of credibility and auditability I relied on two main approaches to data collection. On the one hand I emphasized the collections of data from several sources (field notes, records, documents, taped interviews) in order to follow the principle of triangulation and corroboration of critical observations and actions. Second, since all of the program evaluations involved several other persons working in research capacities I made the research negotiations the topic of unstructured interviews with team members and selected independent observers of the research process.

There was an extensive variety of data sources, ranging from personal field notes and records to public policy documents. These included my own task/time record which included a day-by-day description of the specific research tasks, who was involved in the tasks, and how long it took to accomplish them. This record served as a detailed activity log as well as the basis for keeping track of research costs and billing to clients. Also my field notes included observations and conscious attempts to analyze the interaction of the main actors engaged in the evaluation process. Several portions of the notes set out details of the timing of activities, describe the settings of specific types of interaction, and the parties involved in specific organizational maneuvers.

There are more than one hundred hours of taped interviews with the major actors in each of the case studies; most of these are from the Strathcona case study and the case study of the LSS Schools Program. For the other two case studies there are notes from structured interviews with the major actors (10 from the NIP/RRAP evaluation and 6 from the JFSA study).
In addition there are several hours of taped records of the meetings of the working teams in the four case study setting. These are augmented by records of all correspondence among the major actors in each study, program files and financial records, and notes from unstructured discussions with selected policymakers. There are a substantial number (over 50) of research and policy documents related to the case studies, case study areas/target groups, and the various funding organizations which provided background material for analyzing the structural context of the case studies.

Although memory becomes a factor in this type of secondary analysis the range of data collected while the program evaluations were in process has provided the resources to "cross-examine" my experience of the past. In addition, during the final year of writing and re-writing the dissertation I did focused interviews with all but one member of the various study teams and several of the main actors which dealt generally with the final construction of the research document and its reception and subsequent use by the clients/consumers who had funded the research. Again, this was done in an effort to expand the 'memory-base', and collect the reports and interpretations of others who had been engaged in the program evaluation process. In concluding this section on methodology it is important for the reader to recognize that I have used the negotiated order paradigm in an heuristic fashion and do not claim to be producing descriptive ethnography in the symbolic interactionist tradition. My methodological emphasis has been centered on a strategy of triangulation, relying on data from documents, records, tapes and interviews as well as the notes from field observation.
Models of Negotiated Knowledge Production and Use

The models of negotiated knowledge production and use developed in Chapters 4 and 5, and described more fully in Chapter 6 are designed to be preliminary formulations to be explored more systematically in further research. In that final chapter I present four possible models, suggested in the course of the case study analyses, which I have labelled experimental, managerial, collaborative and transformative. These labels have been given partly as quick descriptions, but also as representing some important aspect of the dominant area of negotiation in the evaluation process. For example, in the experimental model of knowledge production and use, the program evaluator takes the familiar consensual approach to overall research strategy, sharing with the major actors (usually limited to representatives of the funding agency and management of the program) a positivist model of evaluation research and reliance on objective measures of program impact. All the evaluation research is designed to be accomplished by evaluators whose work base is external to the program under scrutiny. High technical standards are important for the evaluation researchers and close adherence to experimental design are key goals for program managers and staff. Therefore, the resulting research report is constructed in a negotiation context where conflict is minimized and objective, numerical evaluation findings are given great credence. The final evaluation document is presented to the funding agency for interpretation and possible future use in the policy process. There is minimal evaluator input into subsequent phases of policy development and program modification.

Another pattern of knowledge production and use of policy research can be defined as the managerial model. This is the model frequently employed in managerial reviews of ongoing programs or program components where the evaluation research is done by "in-house" researchers and the major actors are all members of one organizational context (e.g. Ministry of Education, Central
Mortgage and Housing). Here too the evaluator(s) employ a consensual research strategy, where consensus implies acceptance of the theory and practice of "good management" and is the main focus of program assessment. There is an explicit emphasis on the discovery of the day-to-day facts of program operation, careful recording of individual accomplishments and failures, and a steady reliance on whatever research models and methods have been followed by staff in assembling the data required by program monitoring procedures. This does not always mean quantification of everything under scrutiny, but it does mean accepting the categories and labels given to data in the ongoing process of program management, e.g., substandard housing, unemployment rates, schizophrenic patients, persons ineligible for program benefits, etc. There is an emphasis on statistical or observational conclusions which are not "counter-intuitive" to the accepted wisdom of the Ministry/agency, and the resulting evaluation report may frequently be used as grounds for firing personnel, eliminating a project/department, or alternatively as a means to cover up problems in program delivery which have been defined deliberately as lying outside the parameters of the management review.

The third pattern developed more fully in Chapter 6 has been called the collaborative model. This model may be used in organizations which are attempting to make gradual changes in their structure and long term operation. Thus, programs which have reached some type of organizational crisis point in terms of work overload/employee turnover/rapidly shifting demands for program service tend to look for a way to evaluate their existing situation and make systematic improvements. In this model there is often an attempt to involve as many actors or representatives of interested parties as the evaluation process can effectively include. The evaluation is located both externally to the program (often through use of specific consultants) and simultaneously as an "in-house" process of self-analysis and internal research. Given that there are conflicting
theories of intervention represented in most reform programs, the evaluation negotiations frequently center on developing a research strategy that is minimally acceptable to all parties. Then the evaluation researchers tend to make use of an array of models and methods which will be approved and understood by different components of the program's staff and service users. The final form of the evaluation report is developed collectively as a documentary review of program changes in process, including minority policy positions and reports of program inadequacies and failures. Although some form of report goes into "document time", most research results are utilized immediately by various parties to the evaluation through a process of continual discussion and feedback. Since this model tends to focus on a form of collaborative effort, small scale conflicts can usually be managed within the negotiation process. However, significant differences in intervention theory and major differences in power or resources among the potential partners can often produce serious obstacles to the achievement of long term organizational changes in program policy and/or operation.

The fourth model of knowledge production and use presented in Chapter 6 has been labelled the transformative model and may be the least common approach to policy research presently in use. In this model program evaluators seek active support and involvement of those actors least likely to be represented in the other models, i.e. the program participants and service users/consumers. The evaluation is specifically designed to be external to the program and at some conceptual distance from the theories and assumptions of program funders and managers. The research strategy tends to range from the contentious approach to a paradigm-challenging stance which calls into question program policy, funding, structure and operation. The choice of research models and methods may vary, but where the analysis relies on a political economy or Marxist approach there may be significant reliance on positivist models and
and an emphasis on objective methodologies. The methodological details may depend to some degree on the parties negotiating and the constituency to be addressed in the evaluation document. Frequently the final form of the evaluation report is developed by the evaluator(s) and other community representatives as a document to provoke critical discussion of program policy, operation and change. It is structured to emphasize the gaps between program theory/goals and the limited objectives of day-to-day program service. Program evaluators and community representatives often plan how research results may be used to educate program clients/consumers and to repudiate the findings of program evaluations which have been done using a managerial or experimental model of knowledge production and use.

These four models do not exhaust the possibilities of using the negotiated order paradigm to explore variations in the policy research process. Furthermore, there are elements of each model which may be used by various parties to the negotiations in any specific program evaluation context. But I have limited my discussion in Chapter 6 to these four because they are reflected most obviously in the case studies in which I have been most extensively involved. In the concluding portion of Chapter 6 I suggest a plan for "sampling the universe" of program evaluation studies to develop additional models of knowledge production and use in the area of policy research.
NOTES TO CHAPTER 1

1. It is important to point out that program evaluation is still "an open employment frontier" for sociologists at a time when the academic marketplace is shrinking quite rapidly. Scrutiny of the "job openings" listed in traditional sources of employment information show that program evaluation represents a major source of non-academic employment for social scientists (Eaton 1980).
CHAPTER 2: CONTEMPORARY MODELS OF PROGRAM EVALUATION: PARADIGMS AND LIMITATIONS

2.1 Introduction

This chapter is intended as an overview of the principal models of research that program evaluators claim are most frequently (and most appropriately) employed in doing their work. It is important to emphasize that discussion of these models, as "ideal types", transcends the enormous variations in scope, setting, and objectives that are characteristic of programs as they operate in the field. This critical review of two major models of evaluation research centers on the field of education where program evaluation has received extensive attention in the development of public policy. The description and analysis of each model includes discussion of its main assumptions and directives regarding resolution of the six theoretical/practical issues of policy research.

Despite marked differences in the context in which program evaluation is practiced, there tends to be minimal variation in the research models implicit in that practice. Analysis of the underlying approach taken in the texts on program evaluation (Rossi and Williams 1972; Weiss 1972a; Cano.1971; Rossi Freeman and Wright 1979) reveals a rather limited set of prescriptions for accomplishing the work of evaluation. Although many well known examples in the literature reflect certain individual preferences (Guttentag and Struening, 1975), there is one plan for evaluation research which has tended to predominate, i.e., a conventional view which derives basically from the positivist paradigm of social science (Rossi and Williams 1972; Boruch and Riecken 1975). In the last decade another view of evaluation research has emerged to challenge the positivists. This second view derives, naturally enough, from an interpretive (sometimes referred to as "anthropological") paradigm of social science (Parlett and Hamilton 1976; R.E. Stake 1974).
It should be noted that the two distinct paradigms, frequently labelled the "quantitative" and "qualitative" orientations to knowledge, have both evolved most extensively within educational research. In an article which reviews and analyzes the relations among educational research paradigms, Rist (1977:43) points out that when we speak of quantitative or qualitative methodologies:

> We are, in the final analysis, speaking of an interrelated set of assumptions about the social world which are philosophical, ideological, and epistemological. They encompass more than simply data gathering techniques ... the selection of a particular methodology is profoundly theoretical, regardless of its relative availability. Research methods represent different means of acting upon the environment. To choose one line of action over and against another is to have foregone others available from a different perspective and orientation. Each method reveals peculiar elements of symbolic reality. And to accentuate one aspect of that reality vs. another is to influence both observations and conclusions (Denzin, 1970:298). All knowledge is social.

What follows in this chapter is a brief critical description of these models of evaluation research as developed in the field of educational programming. It is intended to outline the connections between: (a) their underlying paradigms, and (b) their more commonly discussed "prescriptions" regarding the correct forms of research design and methodology. In addition, I point out some limitations and criticisms of each approach in a format that reflects the basic controversies which take precedence in negotiations between "ideal type" evaluators committed to each of the two models.

2.2 **The Agricultural-Botany Paradigm and Conventional Evaluations**

The dominant paradigm in educational evaluation is the "classical" or "agricultural-botany" approach to knowledge which makes use of a hypothetico-deductive methodology derived from the experimental and psychometric traditions in psychology (Parlett and Hamilton 1976). The label "agricultural-
"botany" is deliberate according to Parlett (1972). Many of the statistical and experimental techniques used in education research were developed originally (e.g. by Fisher) for use in agricultural experimentation. Most commonly this model of evaluation is presented as an assessment of the effectiveness of innovation (e.g., reading program, classroom organization, curriculum) by examining whether or not it has reached certain standards on a predetermined set of criteria. The analogy between students and plant crops becomes rather obvious. That is, students are given pretests (seedlings are weighed or measured) and then submitted to different learning experiences (treatment conditions). After a period of time has elapsed, their attainment (growth or yield) is measured to indicate the relative efficiency of the methods (fertilizers) used. (Parlett and Hamilton 1976). Studies of this kind are designed to produce data of one particular type, i.e., "objective" numerical data that permit the use of statistical analysis. Isolated variables like IQ, social class, test scores, personality profiles, and attitude ratings are codified and processed to indicate the efficiency of new curricula, media or methods (Light and Smith 1970).

This mainstream paradigm of knowledge production in its more 'prescriptive' form as a model of evaluation research dominates much of the massive two volumes (1400 pages) Handbook of Evaluation Research (Guttentag and Struening 1975). In that text and others (Rossi and Williams 1972), evaluation is presented as basically a scientific or technological task. The work of the impartial, highly skilled investigator is to provide a "firm assessment" of a program or institution (Rossi, Freeman and Wright 1979:21). Therefore the evaluator enters the situation after the initial events in the life of a program have taken place, works alongside the program for a time making observations and collecting data, and concludes her efforts in a report to an all-powerful decision maker (usually located outside the program under consideration).
The prescribed format for design and methodology in this model of evaluative research can be conceptualized as a series of eight basic steps (The Stanford Evaluation Consortium 1976):

1. **Proposal for Services**: An agency or individual proposes a new service or changes in existing services.

2. **Demand for Evaluation**: As part of a process of gathering support for a proposal, or as an ordinary management tactic, an evaluation is required. Some person or work team, inside or outside of the operating program, becomes the evaluator (E). Issues concerning continuation, modification, or expansion of program services are placed on the future agenda of a decision-maker (D) who is supposed to use the "objective facts" which will be provided by E.

3. **Program Realization**: Many persons under agency supervision and according to planning guidelines turn the original proposal into a real program. Short or long term approval is received and the evaluator enters the scene just as this stage is completed. The evaluator has had little or no input in the designing of the program itself.

4. **E Formulates Questions**: Once E has been formally hired or assigned to the task of evaluation, she takes the first conventional steps of positivist scientific method --- stating the questions to be answered or hypotheses to be tested. Some writers say that to formulate these questions properly E must construct a conceptual model of program operations and effects (Riecken and Boruch 1974).

5. **Delivery**: The decisions about research strategy must be made at this point in the life of the program. Conceptual factors are central to this work, but E must also take into account practical, political and ethical factors. There are two alternative design strategies:

   a) **E Sets Assignment Rules**: In some settings E may determine who will receive a particular treatment and who will receive either no treatment or a contrasting treatment. She then confines her attention to making sure that subjects are assigned to treatments according to her design. This is social experimentation strictly defined.

   b) **E Records Participation**: Frequently E cannot dictate treatment assignments. For example, she is hired only after some communities have chosen to participate in a new English Language program, or a school system has adopted a new service, or some families (not others) have chosen to participate in a new program for pre-schoolers. In these circumstances most "design" options are restricted. E's task is to collect data on the program as it stands --- she records who enters each treatment group, or who is in each group at
5. b) (Continued)
the point of her arrival, and background data on all par-
ticipants. She is free to speculate about program possi-
bilities not in effect and to collect any data that appears
relevant to such possibilities.

6. E Measures Outcomes: The next step is to measure the impact
or outcome of the program according to the predetermined cri-
teria of effectiveness. Has participation in Project Headstart
increased the reading and vocabulary levels of children in
grades 1 - 3? What are the differences in achievement test
scores among grade 8 students who take math classes and those
who follow a self-pacing program? E may choose to design and/
or administer measurement instruments of her own or may rely
extensively on secondary sources, replications of existing
methods.

7. E Processes the Data: This is a step that requires a high
level of technical expertise. While the design and measure-
ment phases necessitate complex research skills, the data
processing phase is the area where virtuoso skill with computer
software and statistical tools is strongly valued in this
research model.

8. E Reports Her Conclusions to D: The work of evaluation has been
done. If it has been done well, the decision maker is presented
with the evidence she requires to proceed rationally with future
program operation. If the program fails to produce the promised
benefits, to an acceptable level, then it is terminated to save
resources. In general, the program whose benefits have out-
weighed its costs is maintained or expanded.

These eight steps outline a process for evaluative research which is con-
sistent with the theoretical and methodological requirements of the positivist
paradigm. The resulting product is an evaluation document which is presented
as an objective, unbiased assessment of the program under scrutiny. When
handed over to the decision maker(s), the evaluation report enters "document
time" (Smith 1973) where much of what has gone into the making of that evalu-
ation is obliterated. This further enhances the claim of objectivity and
usually makes it impossible for others (e.g. teachers, principals, parents,
members of the media, other social scientists) to discover the assumptions and
interpretive work which went into production of the evaluation document.
Finally, it is difficult to grasp the nature of the management practices of
the program under evaluation, and frequently impossible to obtain any useful
indication of the structural or negotiation context in which the evaluation document was produced. This is the situation one faces in reading a program evaluation report constructed in the "purest" form of the positivist tradition.

Despite these limitations in the presentation of information regarding the context of program implementation and service delivery, this model of program evaluation is often presented as the standard against which all work in the field should be measured. Its strengths are defined in terms of the objectivity of its methodology and the ease with which one can generalize from specific results in a sample of cases to the large universe of similar "treatments" in similar schools and classroom settings. Particularly in the evaluation of heavily funded, large-scale national programs, the policymakers and evaluators tend to depend almost exclusively on this positivist model for making their overall assessment of program impact.

2.3 The Social Anthropology Paradigm and Illuminative Evaluation

Again referring back to the area of educational research we find the clearest attempt to develop alternative approaches to program evaluation. In this context it is important to recognize that we are not dealing with a situation of parity among the various research methodologies and models of evaluation (Rist 1977). In educational research the agricultural-botany paradigm and quantitative methods are dominant. This approach is more widely published, accepted, taught and rewarded because of a continuing belief that this is the type of research that constitutes scientific knowledge. However, conflicting intervention theories in education and increased demands for control by the consumers of educational programs have encouraged the evolution of another model of program evaluation in the field of educational research.
There are variations in details of the alternative view of evaluation research; however, the epistemological source lies in the idea of an interpretive social science and a paradigm of social science which is common to parts of sociology and social anthropology (Fay 1975; Giddens 1976; Parlett and Hamilton 1976). The primary concern of this paradigm is with description and interpretation rather than measurement and prediction. As a model of evaluative research it takes account of the wider contexts in which innovative programs function.

The approach described as "Illuminative Evaluation" (Parlett and Hamilton 1976) provides an example of one of the major alternatives to traditional models of program evaluation. The aims of illuminative evaluation are to study the innovative program in terms of (Parlett and Hamilton 1976):

- How it operates.
- How it is influenced by various school situations in which it is applied.
- What those directly concerned regard as its advantages and disadvantages.
- How students' intellectual tasks and academic experiences are most affected.

In the process of evaluating a program the main purpose is "to discover and document what it is like to be participating in the scheme, whether as teacher or pupil, and, in addition, to discern and discuss the innovation's most significant features, recurring concomitants, and critical processes" (Parlett and Hamilton 1976). The model seeks to address and illuminate a complex array of questions which differ quite considerably from the "treatment" and "impact" questions which are central to the traditional model of evaluative research. It is claimed that (Parlett and Hamilton 1976).

"Research on innovation can be enlightening to the innovator and to the whole academic community by clarifying the processes of education and by helping the innovator and other interested parties to identify those procedures, those elements in the educational effort, which seem to have had desirable results."
Illuminative evaluation is not presented as a standard methodological package but as a general research strategy. It aims to be both adaptable and eclectic. The choice of research tactics follows from decisions in each specific case as to the best available techniques, not from a generalized research doctrine. "The problem defines the methods used, not vice versa" (Parlett and Hamilton, 1976). Given the limitations of different methods, no method is used exclusively or in isolation—different techniques are combined to throw light on common problems and cross-check tentative findings through a process of "triangulation".

This shift in paradigm entails more than an exchange of methodologies; it involves new suppositions, concepts and terminology. At the outset the evaluator is concerned with familiarizing herself with the day-to-day reality of the setting(s) she is studying, focusing in particular on the "instructional system" and the "learning milieu". These two concepts are central to an understanding of illuminative evaluation. The "instructional system" refers to those formalized plans and statements which relate to particular teaching arrangements. Each of these plans/statements constitutes or defines an instructional system and includes such elements as a set of pedagogic assumptions, a new syllabus and details of techniques and equipment. The traditional evaluator builds her study around innovations defined in these formal plans, extracting program goals, objectives and outcomes. From these, in turn, she derives the tests, attitude and skill inventories that she will administer in order to evaluate the "effects" of the "treatment". This technological approach fails to recognize, or ignores, the fact that instructional systems when adopted undergo significant modification by teachers, administrators and students as they interpret and reinterpret the instructional system for their particular setting.
The switch from discussing instructional systems in their abstract form (as in conventional evaluation research) to describing the details of its implementation in specific settings requires the second concept, the "learning milieu". This is the social-psychological and material environment in which students and teachers work together. It represents "a network or nexus of cultural, social, institutional and psychological variables" (Parlett and Hamilton 1976). The configurations of each learning milieu, its diversity and complexity, the interaction of various factors—acknowledging all these is considered an essential prerequisite for the serious study of educational programs in this alternate approach to evaluation. In the course of doing illuminative evaluation the researcher makes no attempt to manipulate, control or eliminate situational variables. She takes the complex scene as given and her main responsibility is to "unravel it, isolate its significant features, delineate cycles of cause and effect, and comprehend relationships between beliefs and practices and between organizational patterns and the responses of individuals" (Parlett and Hamilton 1976).

Characteristically there are three stages in the illuminative evaluation. These steps overlap and are functionally interrelated, but can be presented as conceptually distinct. The investigators observe, inquire further and then seek to explain. The observation phase occupies a central place in illuminative evaluation. The investigator builds up a continuous record of ongoing events, transactional and informal remarks and seeks to organize these data at the source, adding her own interpretative comments. She is present at other events which relate to program operation. The second stage begins with the selection of a number of phenomena, occurrences, or groups of opinions which serve as the topics for more sustained and intensive inquiry. While the first stage is aimed at making the evaluator more "knowledgeable" about program operation, at the second stage she makes her questioning more focused and her observations
more directed, systematic and selective. The third stage involves seeking general principles underlying the organization of the program, locating patterns of cause and effect within its operation, and placing specific findings within a broader explanatory context. Within this three-stage framework an information profile is assembled using data collected from four areas: observation, interviews, questionnaires and tests, and documentary and background sources. The process entails building up an extensive data base and then systematically reducing the breadth of the inquiry in order to give more concentrated attention to the emerging issues. The advantages of this model stem partly from this "progressive focusing" which permits unique and unpredicted phenomena to be given due weight, reduces the problem of data overload, and prevents the accumulation of a mass of unanalyzed material (Parlett and Hamilton 1976:146).

To summarize the major characteristics of illuminative evaluation it is important to stress that it belongs to an anthropological research paradigm in contrast to the agricultural-botany paradigm which underlies most conventional program evaluation. In illuminative evaluation the attempted measurement of "educational products" is abandoned for intensive study of the program as a whole including its evolution, operations, achievements and difficulties. The evaluator concentrates on "process" within the learning milieu, rather than on "outcomes" derived from a specification of the instructional system. The methodological strategies are all combined to help "illuminate" problems, issues, and significant program features (Parlett and Hamilton 1976:141). In a general sense all those who articulate the need for a revision of the conventional model of program evaluation, especially in its prescriptive format, move toward a more extended version of the field and focus of such research.

The Stanford Evaluation Consortium in proposing their outline of an alternative model (1976) emphasize certain general features which amplify Parlett
and Hamilton's work. Briefly, the Stanford model\textsuperscript{10} emphasizes the following steps or stages in the life of a program:

1. **E Undertakes Planning Studies**: The evaluator can constructively enter the picture earlier, be active even before the decision to launch a program, helping to define the needs to be met and to understand the potential client population. Planning studies require close collaboration between evaluator and program developer. The plan becomes their joint product, often developed in discussion with members of the client community and officials outside the immediate system.\textsuperscript{11}

2. **E Formulates Questions**: In contrast with the conventional model, research questions are expected to emerge as the study progresses. In practice, the program goals and values of participants change during the evaluation of a program.\textsuperscript{12} Variables and issues emerge as salient only after the team is in the field. E may find that late in the evaluation period formal measures may have to be developed for phenomena whose importance has just been recognized. Thus E need not limit her concerns to objectives stated in advance; she can function as a naturalistic observer whose inquiries grow out of her observations. She need not concentrate only on outcomes, but study just what was delivered and how people interacted during the process.

3. **E Reflects and Discusses**: According to the conventional model all data is stored up for use in an ultimate report to the decision-maker. In the alternative model E can use the data as they are collected from the field and given preliminary interpretation. This information is helpful to program managers and the evaluation staff to improve what both are doing. Change is inevitable. Once this is acknowledged in the evaluation model, the same logic that warrants E's engagement at the planning stage warrants her engagement in program modifications. This involvement is underscored by the recognition that systems are rarely influenced by reports received in the mail.

4. **E Reports Her Current Picture to Outsiders**: Findings are fed into an adversary process, operating programs evolve and change makes evaluation results outmoded after a time. E can remain as part of the picture by returning to involvement in the planning mode of stage (1). She explores the plausible further directions of program development, checking to see what might appeal to various segments of the decision-making community and making predictions about program alternatives on the basis of her accumulated knowledge.

These four stages outline the initial cycle of a relatively integrated process of planning, implementation, evaluation and change. However, there are two basic problems inherent in this model of program evaluation. First, in this
model we must assume a negotiation context in which all disagreement can ultimately be resolved through an adversary arrangement among actors with equal power. Second, there is little or no recognition given to the structural context of the program being evaluated. The structure and policies of the educational system (or medical, housing, legal systems, etc.) are not given any obvious attention in this interpretive model. This lack of recognition of intrinsic economic/political conflicts and a history of bureaucratic struggles may frequently put the interpretive evaluator in a position of much greater vulnerability than the positivist who claims the protection of her objective methodology and value-free professionalism. There are serious obstacles to the construction and use of illuminative evaluations in structural contexts where long term differences exist regarding theories of intervention (e.g. segregation or mainstreaming of handicapped children) and/or where there is a significant imbalance of power and resources available to those actors who fund, deliver, and receive service (e.g. the traditional pattern in schools serving Native people). In these contexts evaluation reports which rely on the purest form of the interpretive model are frequently criticized as being "political" or "idealistic" by policymakers and school superintendents who were themselves educated in the tradition of objective knowledge and impact measurement characteristic of the positivist paradigm of program evaluation.

2.4 Limitations and Criticisms from Within Each Paradigm

The following is my construct of an unresolvable series of criticisms and limitations presented from the perspective of supporters of the two paradigms. It would, obviously, be dominated by the unreflective silence of the positivist and the increasing reflexivity of the interpretive evaluator. As might be anticipated the interpretive evaluator would stress the gap between the
positivist model and the uncontrolled nature of the settings in which it is applied, while the positivist would tend to emphasize the subjectivity and potential unreliability of interpretive models and findings. However, neither one is likely to bring forward criticism which relates to the overall political and economic context in which programs are planned, implemented, and constantly re-shaped.

The Interpretive Critique of the Positivist Model

The strengths of the conventional model of program evaluation research (which ideally takes the form of the classical controlled experiment) are based on the assurance that the study design offers a conceptual framework for testing both the internal and external validity of the evaluation research. Advocates would argue that careful research design and controlled experimentation are the only way in which evaluators can determine the effects of a program accurately and generalize from the results of one program to others in similar settings (Boruch and Riecken 1975). Randomization and the use of control groups provide the critical tools for measuring the impact of the pre-determined independent/dependent variables and outline objective, statistically reliable procedures for determining the population to which policy planners can generalize the evaluators' results. The positivist evaluator can rely on a research model which has a long, respectable tradition in the physical and biological sciences. It is the transferability of that model to the social sciences which would likely be the focus of the interpretive evaluators' initial critique.

There are several other types of problems which are central in the interpretive critique. These begin with the assumptions which characterize the agricultural-botany paradigm. It is presumed that the researcher is able to: (1) specify and operationalize all relevant parameters of the problem and (2) randomize treatments using large samples. One difficulty with these
assumptions is that they entail a major data collecting exercise (expensive in time and resources) and run counter to the need, widely acknowledged, for evaluation before large-scale application, rather than after it. In addition, assumption (2), which involves strict treatment control, is rarely followed in practice. To attempt to simulate laboratory conditions by "manipulating" personnel (e.g., health care workers, teachers, welfare intake people) is not only dubious on ethical grounds, but leads to gross administrative and personal inconvenience (Parlett and Hamilton 1976:143). Therefore, even if a situation could be so unnervingly controlled, the artificiality of the results would make the exercise irrelevant to the actual full-scale operation of the program. It is rare that "tidy" results can be generalized to "untidy" reality.

In addition, such before and after research designs assume that innovative programs will undergo little or no change during the study period. The evaluation is to appraise a particular program, not a shifting target. Freeman (1964:194) has expressed this ideal rather graphically, stressing that "once (the impact model) is formulated, (the researcher) must continue to remain within the environment, like a snarling watchdog ready to fight alterations in program and procedures that would render his evaluation efforts useless". This built-in premise is rarely upheld in practice. However, this assumption of no change remains fundamental to the design and researchers are constrained from adapting to the changed circumstances that arise as part of program delivery. Fidelity to the original design may actually have a negative effect on the program by discouraging new developments and redefinition of aspects of program delivery. In extreme form, this assumption virtually negates the possibility of doing effective longitudinal studies.

The methods used in traditional evaluations impose artificial and arbitrary restrictions on the scope of the study. Concentration on seeking quantitative information by objective means can lead to neglect of other data,
perhaps more salient to program operation, but which are disregarded as being "subjective", "anecdotal", or "impressionistic". An evaluator committed to the interpretive model would stress the necessity for ethnographic data, and emphasize their use in order to satisfactorily explain the findings, assess their importance, and place them in context. On the other hand, positivist evaluators would tend to argue that the facts can speak for themselves. An extension of this problem is that the conventional evaluation model tends to be insensitive to local differences and unusual effects. Since it employs large samples and is designed to seek statistical generalizations, any atypical results are seldom studied in detail. Despite the significance of differential patterns of response to an innovative program, or the possible importance to individuals and groups concerned with program operation, users of the traditional model tend to downplay variation and stress the overall figures in the final report.

The decision-maker in this model is presumed to be rational in the sense of emphasizing value for money. According to this stereotype, when rationality takes over from political considerations resources will be spent where greatest benefits can realistically be expected. Each decision is seen as an isolated activity concerning a distinct separate program. It is presumed that the proper course of action will be unmistakable, once dependable facts are assembled to make that "go/no go" decision. However, the interpretive evaluator would argue that a social agenda almost never calls for a choice between fixed alternatives; "go/no go" decisions are a rarity; newer, more acceptable alternatives are constantly under consideration; and concepts of efficiency and costs and benefits have no clearcut, agreed upon definition among all relevant decision-makers. Therefore, a study which is limited to the original, narrowly defined questions ignores much that is important in the decision-making process.
Finally, this type of evaluation often fails to articulate the varied concerns and questions of program participants, sponsors and other interested parties. The classical evaluators believe that they will produce an "objective truth" that is value free and consequently equally relevant to all parties. Their studies rarely acknowledge the diversity of questions posed by different interest groups, nor take into account the impact of cultural, political and economic characteristics of the setting on the program and on the evaluation.

The interpretive evaluator would be likely to conclude her analysis by pointing out that applying the agricultural-botany paradigm and a traditional positivist model of evaluation research to innovative, reform-oriented social programs is generally an awkward, inadequate and perhaps misleading procedure. It is likely that she would not argue against any use of experimental designs, cost/benefit analysis or conventional survey methods. Rather, for the reasons suggested, they are inappropriate/ineffective as the only form of research method for evaluating the phenomena of social change. She would stress that the positivist approach to evaluation falls short of its own tacit claims to be controlled, exact and unambiguous to the point where rarely, if ever, can it meet most of the basic design requirements of the model. Social reform programs involve a congeries of internal and external influences which can never be controlled in the same way that the researcher controls "treatments" in the lab setting. So the traditional evaluator either ignores or trivializes these influences. She is restrained by the dictates of her paradigm to seek generalized findings, and by the requirements of her model of the "good evaluator" to present these findings in completed form to the decision-maker who alone will act on them. One profound consequence of the positivist approach is that it diverts everyone's attention away from a questioning of evaluation practice toward an automatic acceptance of bureaucratic control over information and a centralized decision-making process.
The Positivist Critique of the Interpretive Model

Given the reflexive nature of the interpretive paradigm, basic criticisms and limitations of this alternative model of evaluative research have been addressed, to a certain extent, by the evaluators themselves. Characteristically, the problems and possibilities of illuminative evaluations are taken up by the authors Parlett and Hamilton in terms of questions regarding its underlying approach to knowledge and issues concerning design and scope of investigation. They have incorporated into their approach a strategy for confronting selected criticisms and formulating a response designed to counter the critique from the positivist perspective.

First and foremost among the questions is concern over the subjective nature of the approach. Skeptics ask whether "personal interpretation" can be scientific, i.e., is not the collection, analysis and reporting of data entirely at the discretion of the researchers themselves? (Parlett and Hamilton:159) The authors' response is that "behind such questions lies a basic but erroneous assumption—that forms of research exist which are immune to prejudice, experimenter bias, and human error. This is not so!" (p. 150). Epistemologically speaking any research study requires skilled human judgements and is, in that sense, vulnerable to human input at every stage. Even in evaluation studies that handle automatically processed numerical data there are several varieties of human selection and interpretation processes at work. Nevertheless, the methods used in illuminative evaluation still raise the possibility of gross bias on the part of the researcher. Parlett and Hamilton suggest that several strategies can be used to cross-check and document all critical research processes: theoretical principles and methodological ground rules discussed and made explicit; criteria for selecting/rejecting areas of investigation spelled out; and evidence presented in such a way that others can judge its quality. But even with such precautions the subjective element remains. It is inevitable.
What is seen by the positivist as a limitation (subjectivity, bias) is viewed by the interpretive researcher as an asset; a resource that is present but overlooked by the positivist. But the inversion does not remove the criticism. For any positivist the essential question of reliability always remains and she is forced to ask: How do we know that your findings are right? The interpretive evaluator is not likely to answer that question directly, but would argue that the "correctness" of her findings is contingent on the extent to which her report readers and users of her recommendations share the values and assumptions of her research paradigm.

Furthermore, an interpretive evaluator would emphasize that there is a crucial shift when one abandons the agricultural-botany paradigm, i.e., the role of the researcher is necessarily redefined and consequently her relationship to the array of actors and data sources in the research process is redefined. Thus, in theory and in practice the use of interpretative human insight and skills is encouraged rather than discouraged. Thus, the illuminative evaluator joins a diverse group of specialists (e.g. psychiatrists, social anthropologists, historians) whose job it is to weigh and sift a complex array of human evidence and draw conclusions from it (Parlett and Hamilton 1976).^7

There is another related criticism which focuses on the researcher, i.e., doesn't her presence have an effect on the conduct and progress of the program she is studying? The response to this is that she does have an effect, and indeed any form of data collection creates disturbance (the Hawthorn effect). She recognizes this and attempts to be unobtrusive (without secrecy), supportive (without collusion), and non-doctrinaire (while still being sympathetic). Furthermore, while the conventional model stresses the technical and intellectual capability of the researcher, in this alternative model such capability must be augmented by interpersonal skills. Tact and a sense of responsibility are essential throughout the evaluation, but are most essential at the
report stage when difficult decisions must be made which balance out full reporting with safeguarding of individual privacy. Parlett and Hamilton comment that "the price of achieving the richer, more informative data of illuminative evaluation is the greatly increased attention that must be paid to the evaluator's professional standards and behaviour" (Parlett and Hamilton 1976). Nevertheless the positivist evaluator is likely to remain skeptical regarding a research methodology which is so intimately involved in the ongoing activities of the program under scrutiny. Questions of bias and interference are the risks that interpretive researchers cannot fully overcome in their encounters with others who follow a positivist model.

The second set of questions facing illuminative evaluation concerns the scope of the investigation. In particular: Can it be applied to innovations that are being widely implemented? Are the results and analyses generalizable to other situations? The authors suggest a general strategy for applying illuminative evaluation on a wider scale. This includes intensive study of a small sample of schools, selective study in a more expanded sample and finally, short visits (in the last resort mailed questionnaires) to the remainder of the institutions. In addition, they offer another way in which the study of specific situations can yield more generally applicable insights with either large or small-scale evaluations. This involves a process of descriptive, rather than inferential generalization. That is, learning milieus (despite their diversity) share many characteristics, and, as such, face habitual difficulties and provoke familiar reactions. There is a wide range of overlapping social and behavioural phenomena that accompany teaching, learning and innovation. Despite wide acknowledgement of these phenomena, few have been pinpointed, adequately described, or defined accurately. The collation of abstracted summaries, efforts to define shared terminology, and development of insightful concepts—the contribution of illuminative evaluation—can serve as
aids to communication and facilitate theory building. These would provide the possibility for moving from the particular to the universal, the theory construction procedure which characterizes this alternate model of evaluative research.\footnote{19}

2.5 Implications

In this chapter I have undertaken a critical review of these major models of evaluative research in the field of education for three reasons. First, I wanted to focus on the strengths and limitations of each of these models as "ideal types" intended to guide the work of program evaluators in the field. Second, I wished to outline briefly the controversial questions that might emerge in a confrontation between program evaluators who use these models in their own research work. And third, it became possible then to describe and analyze these models in terms of their main assumptions and directives regarding resolution of the six issues which I consider central problems of policy research.

In examining the positivist model of evaluation research we find that the only actors whose interests and responsibilities receive direct attention are the decision-maker and the evaluator. The evaluation work may be located inside or outside the operating program, but no recognition is given to the differential significance of that organizational placement. There is a tacit understanding that the general research strategy will be consensual, and the specific model and methods will be as close to the experimental ideal as possible. There is minimal analysis of program delivery included in the evaluation since this model is designed to measure primarily program impact or outcomes. The fact making and report construction process is relatively straightforward, relying
heavily on the presentation of statistical results and the equating of statistical significance with social importance.

Looking at the interpretive model we discover that there is an effort to delineate and include a wide variety of actors whose interests are likely to be affected by the evaluation process and results. The program evaluation work is accomplished both within and outside the ongoing program in order to maintain contact and involvement of key actors. Although illuminative evaluation is presented as a general research strategy, analysis of that strategy in specific research settings suggests that the model tends to be limited to a consensual approach, with some elements of a contentious approach where key adversaries are equally powerful and committed to the program. Since this model aims to be both adaptable and eclectic there is often a range of methods and techniques employed in the evaluation process in order to build up an extensive data base and focus attention on emerging issues. The emphasis is on "process" rather than "outcomes" and this is reflected in the presentation of research results and construction of the research document. Major actors have early access to research data and interpretations, while factmaking and draft production are open to constant discussion. The evaluation report entered into "document time" may already be outmoded because the program has evolved on the basis of initial research results. Meanwhile, the evaluator may continue to be involved in the research process by looking at further program alternatives and/or returning to the planning phase of program development and change.

In comparing the two models it becomes evident that they differ on virtually all of the six issues of policy research. Thus whatever variations may be used in a specific program evaluation study, there will be a tendency for the positivist evaluator to focus on abstract features of program output/impact which are relatively simple to measure, while the interpretive evaluator
will emphasize the processual components of program operation and service delivery as they can be ethnographically discerned through observation and interviews with a wide range of major actors in the evaluation process.

There are several implications to be drawn from this comparative analysis. First, exponents of each approach are cognizant of the strengths of their own model, but often find it difficult to recognize: (a) the limitations of their preferred research model, and (b) the complementarity of the two models for comprehensive program evaluation studies. Looking at these research models as frameworks for constructing reality, evaluators have to more accurately assess the risks and sacrifices entailed in producing and using knowledge gained from different "realities". Under what circumstances is it appropriate to rely on the enriched, detailed accounts of a complex program setting and heterogeneous grouping of clients/participants? When does it become essential to employ "scientific methods" which provide extensive summary data and schematic reliability in describing the overall reality of a "macro-system"? These are the questions which program evaluators must ask themselves in the course of the research design process, during their time in the field, and in construction of the evaluation document.

Second, evaluators seldom recognize the extent to which both models minimize the significance of the larger political/economic structures in the context of policy research generally, and program evaluation studies in particular. The research presented in Chapter 3 regarding differential patterns of research utilization (Caplan 1977) suggests that neither model recognizes the importance of the relationship between research producers and those powerful policymakers who have control over long-term program planning and change. The overall research strategy, systematic inclusion or exclusion of particular actors, and the careful planning for research utilization are all issues which become significant in the long term management process in government programs. The
management ideology and information processing approach of policymakers and program funders often determines the use or neglect of an evaluation study, not the details of the research design and methodology.

Finally, this critical review of the two models suggests the need for evaluators to expand their repertoire of possible models for knowledge production and use. The expanded repertoire does not necessarily require a search for new designs and methods, but careful scrutiny of the empirical settings in which program evaluations are accomplished and a more comprehensive analysis of the structural contexts in which they are requested, funded, supervised, and used.
1. Rist begins by sketching in broad strokes what Gouldner (1970) refers to as "domain assumptions" behind the two research paradigms in education. He then proceeds to assess these two methodological orientations in terms of three polarities: (1) reliability v. validity, (2) objectivity v. subjectivity, and (3) component v. holistic analysis. Rist considers that these three provide "a sufficient map upon which to chart the convergences and divergences of the two research paradigms in question" (1977:44). He stresses that "there is no omnibus strategy for our study of causality" and concludes that there has been a shift from "disdain to detente" among supporters of each paradigm. But he warns that it is not appropriate to think in the near future of there being a "grand synthesis" of quantitative and qualitative methodologies (p. 47).

2. The problems and limitations of predetermined criteria are evident in the controversy over the success or failure of Project HeadStart. For a detailed analysis see Lois-Ellen Datta (1976).

3. For a more flexible version of this model see Carol H. Weiss (1972). Her prescription for assessing program effectiveness has a strong positivist base but calls for more qualitative data and is cognizant of the impact of social context and political conflict inherent in the process of evaluation research, as well as issues of funding. The basic theme of the book is that "evaluation uses the methods and tools of social research but applies them in an action context that is intrinsically inhospitable to them" (p. vii).

4. Evaluators trained in traditional research skills have relied heavily on tests of significance. Frequently they have been guilty of equating statistical significance with social importance. (Stanford Evaluation Consortium 1976:209).


6. The Stanford Evaluation Consortium offers another example of an alternative approach in Cronbach et al., Toward Reform of Program Evaluation. That work includes their 95 theses and an extensive critique of the conventional model along with recommendations for improving the evaluation enterprise. It is more a survey of problems and proposals for change within evaluation than a statement of an alternative paradigm.

7. See E.J. Webb et al., Unobtrusive Measures and the discussion in A. Cicourel, A Study of Argentine Fertility.

8. These include: constraints on the organization of teaching in schools; operating assumptions held by faculty, individual teacher's characteristics and student perspectives and preoccupations. Each class/course has a "unique pattern of circumstances, pressures, customs, opinions and work styles which suffuse the teaching and learning that occurs there" (p. 145).
9. This process tends to run parallel to Glaser and Strauss' notion of the discovery of "grounded theory". See Glaser and Strauss (1967).


11. Members of the Consortium point out that in some sense the notion that "the decision-maker" is served by the evaluator is a fiction and perhaps one that has outlived its usefulness. "Numerous persons within the system, in numerous roles, engage in reshaping a program from day to day. Others play a part in the political, legislative and bureaucratic processes that reshape the program from outside" (1977:212). They note that Carol Weiss (Struening and Guttentag 1975:1:17) is one of several writers who warns that there are multiple decision-makers, that coalitions develop and that trade-offs with interest groups, professional guilds and salient publics all play a role. Decisions become more a matter of fine tuning than go/no go in the face of such cross-pressures. For a detailed analysis of this process see Carol Weiss (1977).

12. This notion has had little influence on those government officials who, in contracting for an evaluation, set up a rigid protocol for obtaining specific answers to a limited number of questions. Not all officials operate in these terms, but many sponsors share their evaluators' positivist models of knowledge production.

13. See, for example, criteria for participation in Residential Rehabilitation Component of the SRP (Levitan and Hurwitz 1977).

14. Wacaster and Firestone (1978) present an important analysis of the results of a three year field evaluation of planned change in ten rural school districts in the United States. See 'The Promise and Problems of Long-Term Continuous Fieldwork'.

15. Local differences in the pattern and scope of community participation in the NIP program were crucial to obtaining an accurate assessment of the program throughout British Columbia and in the rest of Canada. See CPAC and Ptarmigan 1977.

16. Brian Fay points out the profound epistemological difficulties entailed in the use of "efficiency" as a criterion by those who advocate a positivist conception of policy science (1976:50).

17. These insights and skills are encouraged for managers and executives in recent work on the development of democratic management in commercial and industrial settings. See, for example, the Industrial studies in Herbst (1976) and Rowbottom (1977).

18. This issue has become critical in large scale studies. For a paper which documents this tension between the accountability for public funds and the assurances of confidentiality that can be given realistically to informants by researchers, See Trend (1980 a).

19. There are many parallels between the approach taken by Parlett and Hamilton and the work of Glaser and Strauss (1967) on the construction of grounded theory. As I have indicated in Chapter 1, and develop further in Chapter 3, Strauss' latest work on a paradigm for negotiated order builds inductively from an interpretive perspective but extends more comprehensively into problems of social structure and historical origins.
3.1 Introduction

In this chapter my focus shifts to the problem of constructing a theoretical framework for the analysis of program evaluation research which extends beyond the limits of the models presented in Chapter 2. After having reviewed two models of how program evaluation should be accomplished, I now direct my attention to developing a theoretical perspective which encourages analysis of program evaluation research studies as they are accomplished in specific organizational contexts in the field. The overall framework presented here relies on the theoretical approach of symbolic interaction, and more specifically on the negotiated order paradigm developed by Anselm Strauss.

From a perspective within the sociology of knowledge, it seems particularly suitable to employ theoretical constructs derived from symbolic interactionist thought to analyze the evaluation of social reform programs. The philosophy of symbolic interaction, as presented by Mead and further developed among social theorists of the Chicago School (Vaughan and Reynolds 1970), clearly reflected its origins in American society of the early 20th century. There was an emphasis on social change occurring through natural processes, the sense of fluidity, and accent on flexible interpersonal relationships in society (Shaskolsky 1970). The symbolic interactionist tradition sought to provide American social theory with a balance between individualism and determinism in the analysis of society. Substantial attention was paid to the communication processes through which people transform themselves and their environment. The topics associated with the symbolic interactionist perspective included: interaction processes, socialization, the establishment and maintenance of meaning, systematic symbolization, as well as social roles, identities, and the reflexive self, which are products of human association (Maines 1977:235).
However, in recent years a growing emphasis on the topics of social organization and social structure, as reflected in contemporary American society, has encouraged a radical transformation in symbolic interactionist thought. Through empirical studies focused on the processes of negotiation in society "symbolic interactionists have considered a wide range of substantive and theoretical issues falling under that general topic (social organization and social structure), including those of power, authority, formal organizational structure, legitimation, autonomy, national institutions, and economic markets" (Maines 1977:253). This transformation of symbolic interactionist theory is particularly well developed in Strauss' negotiated order paradigm (1978) which offers a theory of social organization that attempts to include both macro (structural context) and micro (negotiation context, sub-processes of negotiation) levels of analysis. He has constructed a paradigm which directs our attention to the processes of negotiation which occur at several levels in organizations and argues that use of the paradigm encourages researchers to trace the complex relationships between structural context and negotiation context in a wide variety of settings. In summarizing these relationships Strauss indicates that the various structural conditions of the settings affect: the actions of the negotiating parties; the aims they pursue through negotiation and alternative modes of action; their tactics through negotiation; and, undoubtedly, the outcomes of the negotiations themselves, which in turn may affect not only future courses of action but also the social settings themselves. Consequently there is an iterative process where one can, and should move both ways through this particular paradigm.

In the subsequent sections of this chapter I present, first of all, an overview of Strauss' work on negotiations. It is essential to outline more fully the central concepts and implications of the paradigm for empirical research. Second, using various aspects of Strauss' theoretical argument, as
well as other empirical and theoretical sources, I construct a framework for
the analysis of program evaluation studies which: (a) specifies various factors
in the structural context of policy research generally, and program evaluation
in particular, and (b) presents the six theoretical/practical issues of policy
research as basic areas of negotiation in the development of program evaluation
studies. This will become the framework used in the description and analysis
of the four program evaluation case studies discussed in Chapters 4 and 5.

3.2 Overview of Strauss on Negotiations

The term NEGOTIATION is extremely broad and imprecise if we want to think
of it as a sociological concept. Nevertheless, it has been developed in the
literature of both the social interactionists and the exchange theorists.
Maines (1977) locates the origins of the concern with negotiation in the urban
and occupational studies carried out at the University of Chicago. Since Strauss
has long been connected with the ongoing development of this perspective, it is
important to outline, briefly, the evolution of the negotiated order paradigm
from his point of view (see also Maines 1977).

Negotiation and Social Order

In his most recent book, Negotiations: Varieties, Contexts, Practices
and Social Order (1978) Strauss directs attention to negotiation in its relation-
ship with social order. He points out that when individuals or groups or organ-
izations of any size work together to get things done agreement is required
about such matters as what, how, when, where, and how much of anything will
be done. Strauss emphasizes at the beginning of his discussion that negotiation
is certainly not the only means for getting things accomplished, or for working
toward collective or individualistic ends. He suggests a variety of alternatives
that are available. These include: persuading, educating, manipulating, appealing to the rules or to authority, and coercion. If we look at any society or organization we ordinarily find a mix of all these processes in varying proportions at different times. However, he is particularly interested in negotiation, and his basic question concerns the idea that we must begin to understand under what conditions negotiation is given priority over these other modes, under what conditions it is placed in conjunction with other modes, and what are the consequences for various interested parties.

The conventional assumption is that negotiation is only relevant to certain kinds of activities or certain kinds of areas, e.g., business arrangements or labour relations or international relations. In fact the traditional definitions of negotiation often make reference to those substantive areas. But Strauss feels that that is too narrow a focus and that the other assumption made about negotiation, i.e., that it be confined to rationalistic or efficiency based perspectives, is also too narrow a focus. Strauss wants to develop some kind of theoretical perspective that explicitly links larger structural considerations and more microscopic analysis of negotiation processes. He puts effort into creating a theoretical paradigm with applications to a variety of negotiations. One of the central arguments of that paradigm is that social orders are always in some sense negotiated orders. Therefore we realize that much of Strauss' paradigm will be based on the theory of negotiated order in organizations and it will be focused on the connection between structure and process.

Strauss concludes an overview of the literature on negotiation by suggesting four directives: First, that theorists in negotiation need to look at a much fuller range of negotiation; second, that they need to relate the negotiations to what other modes of action are available to actors who are considering whether to negotiate; third, that they need to look at the negotiations in close relationship to the larger structural contexts within which they take place; and fourth, that
they need to consider the matter of the actors' views or theories of negotiation as these enter and affect the negotiation themselves (1978:11). In reviewing the work of "our favourite theorists and their vocabularies of explanation" Strauss points out the extent to which negotiation processes are assumed, not analyzed. This is true even in the writing of such consensus theorists as Dewey. He remarks, in addition, that certain types of theory cannot easily accommodate the idea that negotiation is very important in analyzing social order. In particular he has in mind the various forms of determinism--biological, social, cultural (1978:13). Ultimately (in his final chapter) Strauss takes up the question: "What degree of freedom versus constraint is implied in one's theories of social order and negotiation?"

**Background to the Negotiated Order Paradigm**

In the late 1950's and early 1960's Strauss and some of his colleagues began to develop the concept of negotiation, along with several others, and coined the term "negotiated order" (Strauss 1978:5). At the time they were doing research on psychiatric personnel in hospitals rather than writing about organizations in general. However, that research suggested that the usual structural and organizational approaches did not give adequate attention to the important phenomenon called negotiation. The main points in that initial discussion are summarized in the more recent work (Strauss 1978:6):

1. That social order was negotiated order.
2. That specific negotiations seemed contingent on specific structural conditions, i.e., who negotiated with whom, when, and about what.
3. That the products of negotiation all had temporal limits.
4. That negotiated order had to be worked at and the basis of concerted action needed to be continually reconstituted.
5. That negotiated order on any given day could be conceived of as the sum total of the organization's rules and policies, along with whatever agreements, understandings, pacts, contracts and other working arrangements were in place at that time.

6. That any changes impinging on the negotiated order called for renegotiation or reappraisal.

7. That the reconstitution of social or organizational order might be conceived of in terms of a complex relationship between the daily negotiation process and a periodic appraisal process.

8. That the complex relationships that exist between the more stable elements of organizational order and the more fleeting working arrangements could be analyzed so that the focus could shift from the more stable to the more quickly changing and vice versa.

In their earlier work, Strauss and his colleagues omitted certain concepts that they have subsequently begun to consider very important, e.g., actors' theories of negotiation, and the sub-processes of negotiation, the conditions, consequences and patterning of negotiation. The new paradigm, developed by Strauss in *Negotiations* (1978), presents an analysis in terms of the concepts of "structural context" and "negotiation context" and examines the options for alternatives to negotiation. Also the new paradigm overtly concerns itself with issues of power, coalition, and politics, ideas that were not handled in the original work. In addition, it examines the temporal features of negotiation, i.e., their timing and sequencing, and the linkage among multiple negotiations, particularly characteristic of large-scale and complex negotiation settings. Therefore it seems apparent that the new paradigm is expanded in order to handle some of the criticisms that have frequently been made of symbolic interactionism and, in particular, the negotiated order perspective (Maines 1977; Strauss 1978).

**Key Concepts in the Paradigm**

 Strauss uses the term "social order" in a very loose sense as referring to "the larger lineaments of groups, organizations, nations, societies, and international orders that yield the structural conditions under which negotiations of
particular kinds are or are not initiated by, or forced, on actors" (1978:12). The definition he develops for "negotiation" is similarly encompassing. That is, "negotiation generally will stand for one of the possible means of 'getting things accomplished' when parties need to deal with each other to get things done. The choice of negotiation as a means is neither fortuitous nor divorced from the social conditions under which it is made" (1978:2).

In this paradigm he distinguishes between the larger structural context and a negotiation context. The structural context bears directly on the negotiation context, but the negotiation context refers more specifically to, what he calls, the structural properties entering very directly as conditions into the course of the negotiation itself (1978:99). The concept of negotiation context is analogous to the use of awareness context in the paper that Glazer and Strauss wrote in 1964 titled "Awareness Contexts and Social Interaction". In that paper they indicate that awareness context is a structural unit, i.e., that it is not a property of one of the standard structural units such as a group or organization, community, role, position. By context they meant that it is a structural unit of an encompassing larger order, larger than the other unit under focus which is the interaction. Thus an awareness context surrounds and affects interaction. Generalizing from that process of the discussion of awareness context, Strauss indicates that there are many specific kinds of negotiation contexts pertaining to interaction among negotiating parties and these differences are related to permutations and combinations of the following properties of any negotiation context. He lists these properties as the following (1978:99):

1. The number of negotiators, their relative experience in negotiating and whom they represent.
2. Whether the negotiations are one shot, repeated, sequential, serial, multiple, or linked.
3. The relative balance of power exhibited by the respective parties in the negotiation itself.
The nature of their respective stakes in the negotiation.

The visibility of the transactions to others, i.e., their overt or covert characteristics.

The number and complexity of the issues negotiated.

The clarity of legitimacy boundaries of the issues negotiated.

The options to avoiding or discontinuing negotiation, i.e., the alternative modes of action perceived available.

These properties of negotiation contexts are not logical constructs, but emerged from the examination of numerous instances of negotiation. Strauss points out for the reader who is particularly interested in method that the properties can be combined in various ways in order to analyze any given case involving negotiations. The chief consideration, both in the article that refers to awareness contexts and in this book on negotiation contexts, is the relevance of the particular set of combinations not the logic in developing specific typologies or analyses of context.

"A steady focus both on structural and negotiation contexts and on their respective properties is likely to increase the possibility that the analysis of specific courses of negotiation will be carefully located 'within' the larger social structure. (No reification intended)" (1978:101). The implication of this statement is that researchers will not overlook the impact of larger social structure on the course of the negotiations within some specific instance. Structural context is larger, more encompassing than negotiation context but in this paradigm the lines of impact can run either way, i.e., changes in structural context may have impact on negotiation context and vice versa. The outcomes of negotiation itself can contribute to changes in negotiation contexts in the future. However, he suggests that they are less likely to affect the structural context except as they are repeated or combined with other negotiations or with other modes of action, so that they might perhaps have a cumulative impact.
Strauss' Use of the Paradigm in Case Study Analysis

The titles of each of Strauss' case study chapters convey the nature of the processes and issues that he considers most important. There are chapters on: continuous working relations in organizations; the interplay of legal and illegal negotiations in the political arena; building cooperative structures; negotiating compromises within social orders; antagonistic negotiations within changing structural contexts; and limits, silent bargains and implicit negotiation. His twelve case studies are paired in six chapters of his text to bring out similar or contrasting features raised by the paradigm. Among the twelve cases the organizational scale varies greatly, the negotiating parties range from individuals and small groups to large nations, the settings vary from face-to-face ones to those that are international, from relatively simple or routine ones to the most complex and unpredictable. Time scales also vary greatly from immediate transactions to those occurring over long periods of time (1978:237). Each case study begins with a description of the negotiations, in terms of the negotiation context. These descriptions include the accompanying types of interactions, i.e., actors, their strategies and tactics, some consequences of negotiations, and what Strauss called the embedded negotiation sub-processes, e.g., making tradeoffs, obtaining kickbacks, paying off debts, and negotiating agreements. In addition, he brings out some of the salient structural properties that bear on the negotiation, the structural context. For example, when he is talking about a corrupt judge, he discusses features of the American judicial system; when he is talking about the structural context of negotiations that occur in mental hospitals he also gives some notion of the properties of American medical care, particularly the subspecialty of psychiatry and the specialization among the caring professions as well as the division of labour in mental hospitals.
Major discussion of the paradigm is located in the first case study chapter which is concerned with "continuous working relations in organizations". I want to outline, briefly, some of the main points in that particular chapter in order to give a sense of how Strauss makes use of his paradigm. Before presenting and analyzing the two case studies he asks what does negotiation look like, what are its special features when people are and must be in fairly continuous working relations. He suggests that obviously no organization can persist unless its members manage somehow or other to come to terms with each other's actions, to work with one another even if antagonistically in order to get things done for, or in, or through the organization. While some relationships among members involve relatively routine actions guided by standard rules or rules of thumb, others may involve conflict, ambiguity, novelty, etc. (1978: 105). He says that such generalized processes as role taking, understanding or the kind of rational processes itemized under decision-making scarcely give the flavour of what really happens when people are working or living in organizations. What happens involves negotiation, but it also involves coercion, threats of coercion, attempts to persuade, appeals to authority, etc. The appearance of these processes and their relationship to what is going on is not fortuitous. So Strauss would argue that "each type of organization, in relation to its structural properties and those of the larger social environment within which it operates, will evidence different negotiation sub-processes that stand in at least somewhat specifiable relationships to alternative modes of action" (1978: 106).

His first case study concerns experimental work in a mental hospital, while the second concerns industrial firms. The mental hospital is characterized by continuous working relationships that to the actors themselves are novel and challenging but often ambiguous, puzzling, risky and frustrating. So there is considerable negotiation over new tasks, new roles, new functions and new and
old jurisdictional terrains. The sub-processes of negotiation that Strauss brings out in that case study involve exploring of legitimate boundaries and staking of territorial claims, since questions of legitimacy are frequent on these wards. In the industrial firms he points out that there is a great deal of characteristic collusion. Covert making of deals is characteristic as individuals attempt to further their jobs, aspirations, careers, and even their non-organized purposes through negotiating covertly as well as overtly. There are other types of sub-processes in both case studies, e.g., mediating and trading-off, but there is little that is characteristic of certain other kinds of negotiation, where for instance, people bargain toward the middle or make settlements as in economic transactions.

A key difference between the two work settings is that on the experimental wards the social order is visibly in the making. The work relationships are evolving before our eyes while the negotiations clearly contribute to all this development. Questions of legitimacy are intimately related to this emergent aspect of the wards whereas in the industrial firms, questions of legitimacy are related more to a relatively durable web of commitments that involves legitimate and illegitimate action. Strauss notes that in both cases the respective researchers were faced with a difficult problem of adequately characterizing the social order of the organizations they were studying (1978:106).

The Strauss Paradigm: Summary and Implications

First of all, in this book negotiation has generally stood for one of the possible means of getting things accomplished when parties need to deal with each other to get those things done. Yet negotiation is not merely one specific human activity or process, because it is important in a particular kind of relationship, e.g., diplomacy, or labour relations, or business transactions; it is of major importance in human affairs since its study brings us to the
heart of studying social order. Strauss argues that a given social order, even the most repressive, would be inconceivable without some forms of negotiation. The implication of that position is that social orders are in some sense always negotiated orders, and he points out the conjunction of negotiation with other processes as mentioned earlier, notably coercion, persuasion, manipulation, etc. Strauss argues that issues of outcome, efficiency, decision-making are secondary to issues pertaining first to how negotiation is related to other modes of action and second to varying kinds of social order.

He also argues that larger structural considerations need to be explicitly linked with microscopic analyses of negotiation processes; negotiations always take place in social settings. We note this stress because this is in response to many of the kinds of criticisms that have come, particularly from the Marxists, regarding the inadequacy or inappropriateness of symbolic interactionist perspectives to deal with the larger social structural elements in society. Clearly Strauss and his colleagues are arguing that the symbolic interaction perspective not only allows for analysis of stable structural elements in social order but is particularly adapted to relating the macroscopic and the microscopic levels. Generally, his argument implies that either a general theory of negotiation should be developed, or at least as he has done, a paradigm should be offered to other researchers to qualify or amend or add to so that a theory would be developed.

He discusses further uses of the paradigm, and here the first question might be how can the paradigm be useful to researchers whose main concerns are not with negotiation in general, but either with negotiations in a given substantive area or only with that area itself. First he says the paradigm can sensitize them to possibilities in their data that otherwise might go unnoticed. Thus they can become sensitized to negotiations, actors, interactions, tactics, sub-processes of negotiation and consequences that have been overlooked or left relatively
unanalyzed in the data. Second, grasp of the paradigm can force them to analyze the negotiation aspects of their data. Researchers will ask themselves if they have enough data on negotiation and along various contextual dimensions. And they will be forced to compare the kinds of negotiations being witnessed or interviewed about, as well as the range of negotiations being studied to other kinds that they know about, so that they will think more systematically and intensively about negotiations. This systematic pursuit of negotiation data should force consideration of relationships between the negotiations under study and the major substantive matters under study.

A third use of the paradigm is that it can help researchers discover which contextual processes of negotiation are salient for their specific substantive materials, and a fourth use of the paradigm pertains to its possibilities as a predictive guide. He argues that the more structural conditions one can anticipate the better the predictions about the main features of both the associated negotiations and the variations in their patterning. The kind of paradigm offered here should help researchers in the theoretical sampling procedures outlined in the Glaser and Strauss book on grounded theory (1967). One difficulty about predicting outcomes accurately is that the structural context may change, thus changing the relationships of negotiation to alternate modes of action among which actors may choose, as well as changing the elements of the negotiation context itself.

A fifth potential use of the paradigm and one implicit in the various cases reviewed is that the paradigm can be useful to researchers who are interested in how negotiated order is destroyed or even prevented from forming. He has in mind not the specific structural conditions that prevent successful, linked negotiations, but the strategies and tactics of parties who are not necessarily involved directly in negotiations. He uses as an example the impact of Allende’s government in Chile which he suggests was undoubtedly destroyed by the course of
events or the falling out of various parties to the coalition, but also by an effective combination of various modes of action—use of force, threat, persuasion, manipulation of events, money, various resources, and also undoubtedly covert negotiations among the various parties who wished to see the Allende alliance destroyed (1978:242).

In concluding this discussion of implications, Strauss says that a vital part of the concept of structural context is that of the negotiating parties' theories of negotiation; hence all of his suggestions also direct researchers to gather data on those theories. One of the considerable deficiencies of studies of negotiation to date is precisely that researchers fail to collect data about those theories, or if they do have some data bearing on them, they do not analyze their significance. Theories of negotiation are usually implicit and he refers back to Gouldner's description of managers' actions (1954) when he indicates how their implicit conceptions of negotiation profoundly affect such questions as when, how, about what, with whom, and how much a manager will negotiate. These unexpressed theories may facilitate or constrain the negotiated order, and deserve greater research attention than they have been given in most traditional studies of organizational or social order.

There is one other question which Strauss deals with here which has implications for theory development, i.e., is a general theory of negotiation possible? Before he attempts to answer the question he states that his position on the possibility of formulating general theory is that it relates to the whole question of historically changing social orders. He insists that we must take into account that negotiations take place within changing structural contexts within historically changing social orders and that the features of a specific negotiation context cannot be properly analyzed except in conjunction with a clear specification of the relevant structural context. The changes brought about by the workings of history do not all signify that a general theory of
negotiation is impossible; they only mean that no such theory can ever predict all the future permutations of negotiation context (1978:244).

3.3 The Proposed Analytical Framework for Policy Research Studies

In outlining my research topic in Chapter 1, I stressed the need to develop a wider selection of policy research models than is traditionally presented in the literature, especially the literature on program evaluation research. One phase of this process of model development involves the construction of an analytical framework which will permit me to both extend and break down the program evaluation process as it is accomplished in various field settings. Such a framework will make it possible to trace empirically the connections between selected social/political/economic/organizational aspects of a program and the pattern of knowledge production and use which then characterizes an evaluation of that program. This detailed documentation of both program background and program evaluation is designed to provide much of the "thick description" (Geertz 1973) which is essential to build up an expanded set of models for doing policy research that has greater contextual relevance.

Background to the Framework

The overall analytical framework developed here relies initially on several concepts in the negotiated order paradigm. To begin with, I outline selected characteristics of the STRUCTURAL CONTEXT in which programs evolve and policy research must be accomplished. These include social, political, economic, and organizational factors which are entailed in program planning and implementation. Next, I suggest that the six theoretical/practical issues of policy research be re-interpreted as basic areas of negotiation in the program evaluation process. This permits me to use the properties of NEGOTIATION CONTEXT as categories to
describe and analyze how the basic areas of negotiation are dealt with in specific program evaluation studies.

However, in addition to my reliance on Strauss' work for providing a general theoretical approach, there are several other theoretical and empirical sources which contributed to the development of the analytical framework to be described in this section. There are two main theoretical sources which have suggested the scope of the framework, Martin Rein (1976) and Jacqueline Wiseman (1979). Rein's essays on policy analysis and, in particular, his case study of the uses of social science research in conflicting approaches to health policy and health care delivery suggest the extent to which policy research is embedded in particular approaches to the social issue in question. In that study he outlines the ways in which models of reality (and the categories for description they contain) have served as guides for linking theory, empirical analysis and policy implications in the area of health care and the definition and treatment of illness. The study argues that there are three different frames for linking poverty and ill health which can be found in the literature--resource allocation (insufficient housing, nutrition, and income), personal and social theory (pathology of the poor, stages in life-cycle), and institutional performance (the poor have least access to medical care services). Selection of one or the other of these competing models of reality commits researchers and policymakers alike to the acceptance of a specific research perspective and priorities for action. Thus, different theoretical and policy perspectives shape different research studies and implications for program change.

A second theoretical source on which I have relied in thinking about a framework for analyzing policy research studies is Wiseman's (1979) overview of the problems entailed in formulating a theory of policy intervention in social problems. Working in the tradition of Glaser and Strauss (1967), Wiseman suggests that we refocus "on the processes of social engineering in order to
learn more about the dynamics of policy intervention from inception to completion" (1979:4). She views policy intervention with regard to a designated social problem as a purposive effort to create a situation in which the structures and processes of social life are deflected in some way by an attempt at change. To fully comprehend these activities requires analyses which are "able to capture the PROCESSES OF CHANGE set in motion by policy intervention, instead of standard before-and-after comparisons." Those who are involved in doing such analyses must be willing to borrow or create concepts which are directly applicable and deliberately aimed at the change process. She notes that in analyzing important aspects of the inauguration of policy intervention it is essential that we include an understanding of (1979:11): how the "facts" about the problem are interpreted and used; what appear to be the goals of policy regarding the problem; how the establishment of either detailed or vague rules affects day-to-day policy implementation; and how money, time, sentiment and skill are resources for and constraints on the various groups involved. Gaining an effective grasp of these components of the policy intervention process should precede our analysis of the series of situations that new policy puts into motion.

Wiseman stresses that there are several requirements of a theory designed to yield a more coherent understanding of the process put in motion by policy intervention. The four main requirements of such a framework include:

1. "Mapping the territory": providing sociological documentation of the territory slated for policy intervention; obtaining more sophisticated descriptions of the multiple realities which exist on several levels for all concerned with a problem; this work to be done before policy is instigated.

2. Locating the ongoing processes: these include the communications process and the processes by which people attain and hold power; locating the processes of policy formulation, initiation and implementation.
3. **Analyzing the ongoing processes**: this means analyzing the series of linked situations and actions set in motion by policy intervention; important elements include the mechanics of commitment, forms of reaction/counterreactions to policy, and the outcome of the policy over time.

4. **Selection of a cutoff point**: methodological closure is achieved through selection of an arbitrary endpoint; assessment of intended/unintended consequences of policy intervention.

While these requirements imply the development of an extensive and complex theory grounded in empirical research, her general perspective on the analysis of policy intervention defines theory construction as a "cooperative and incremental process." Obviously, an approach which stresses that data collection and conceptual development begin before policy inception and "follow the career of policy to its ultimate outcomes" presumes that no single researcher could deal with the entire process in extremely complex, large-scale settings.

Taken together the work by Rein (1976) and Wiseman (1979) suggest the directions in which we might move in an attempt to extend the analysis of policy research studies and locate more specifically the processes of policy formulation, initiation and implementation which precede the work of evaluation in various program settings.

As an additional contribution to this framework I will be including references to a series of empirical studies of policy research in specific settings (policy implementation and program evaluation) which have provided particular insights into the difficulties of resolving basic research issues (Caplan 1977; Kleiber and Light 1978; Bogdan and Ksander 1980; Nader 1980; Stern and Rosettis 1980; Fetterman 1981). However my main empirical source has been my own experience doing a variety of policy action/research work in several different settings. These have included: assisting community groups to organize their resources to pressure government for social reform programs; providing sociological documentation of the territory (groups and neighbourhoods) slated for policy intervention; doing program and facility planning; setting up research
designs for program monitoring; and assessing/evaluating a wide range of programs designed to deliver resources and services (Saroff and Levitan 1969; Levitan and Hurwitz 1977a; Goldberg and Levitan 1977; Legal Services Society Schools Program 1980; Enoch and Levitan 1981).

These empirical sources have been incorporated into the framework for analysis along with Dorothy Smith's (1974a, b) critical work on the construction of documentary reality. Her analysis clarifies the process through which documents come to be produced and read in the organizational contexts which depend upon the continual production of such documents in their ongoing operation.

Analyzing Structural Context

There are four main components of structural context which I have selected to develop here and use as categories for description and analysis in the case studies presented in Chapters 4 and 5. These components, which I have referred to as "social", "political", "economic" and "organizational", represent aspects of the overall processes of policy planning, implementation, ongoing operation, assessment and change which characterize the natural history of most social reform programs. They reflect "the more formalized and permanent organizational rules, policies, established conventions and understandings" (Strauss 1978:6) of the policy intervention process.

The SOCIAL COMPONENT of structural context corresponds quite closely with Wiseman's (1979) notion of "mapping the territory". In examining the social component of the policy and program development process the policy researcher focuses on analyzing the process through which a social problem has been defined. This involves: a) looking at the "facts" which have been interpreted and used in reifying a set of social needs as a "social problem", and b) assembling selected sociological documentation of the territory slated for policy intervention through program development and funding. Frequently this process
of analysis entails critical study of historical material and interviews with those representatives of groups who formulated policy positions early in the process of defining the problem in question (Clairmont and Magill 1974).

The POLITICAL COMPONENT of structural context could as easily be referred to as the theoretical legislative component, since here the policy researcher is concerned with analyzing a) those intervention theories which are current regarding the social problem in question, including the political perspectives of supporters and opponents of the various intervention theories, and b) the policy objectives which have been formally expressed in the legislative mandate of the program, including both the scope of program operation, length of operation (if not ongoing), and the limits of responsibility of the Ministry (or other governmental body) which governs that area of social policy. Information on current intervention theories is usually available from the traditional academic sources and also from non-governmental organizations which are particularly concerned with the policy issue. (e.g. labour economists, women's groups and construction unions regarding development of a program for training women in non-traditional, blue collar occupations). Although reading the legislation provides a literal definition of the policy mandate, analysis of the various handbooks for program managers and staff provides access to the specific interpretation given that legislation in the bureaucratic setting of program operation. Interviews with program managers, staff and clients/participants often reveal significantly different interpretations of the meaning and intent of the same rules or program procedures. These are differences which often give the policy researcher clues as to the irreconcilable or non-negotiable issues in the policy intervention process (Handelman and Leyton 1978).

What I have referred to as the ECONOMIC COMPONENT of structural context requires critical examination of: a) the resource endowment of the major actors who are both directly and indirectly involved in policymaking, program
management and service delivery, and the clients/program participants, and b) their strategies for endowment use. This aspect of structural context can be analyzed in terms of the money, time, sentiment and skill available to the representatives of these groups as resources. When any one or more of these resources is unavailable (or in limited supply), this frequently presents a serious constraint likely to affect the entire structural context of the program in question (e.g. underfunding program implementation plans virtually guarantees a negative program evaluation, Fetterman 1981). Where money is limited, actors often use time, sentiment and skill to make up the deficit. But there are distinct limits to this principle of substitution which are documented in many empirical studies of the impact of underfunding on patterns of citizen participation (Marris and Rein 1974; Hessler and Beavert 1982). In certain situations, the resource endowment of the major actors will be relatively well balanced and there may be a fairly high degree of consensus on how these resources should be used (Rowbottom 1977). However, this is not the norm in most large-scale reform programs, and the strategy of the most powerful is likely to prevent additional resources from being made available to the less powerful (Cronbach 1980:94).

The usual data sources for documentation of resource endowment and strategies for use are contained in budgets and funding statements, lists describing staff members and their responsibilities, time budgets and work schedules which allocate tasks, minutes of management meetings, and the files on clients and/or program participants. The basic problems entailed in "studying up" (Nader 1972) often prevent the policy researcher from gaining access to some of these types of information, especially minutes of meetings, and preliminary forms of budget allocations. There is also the question of privacy of information vs. the public's right to know which surfaces here with regard to program correspondence and client/participant data files (Trend 1980). Information on the resources and resource strategies of participants or interested community groups may often come from the groups themselves, particularly when they lack adequate funding
or time and skills to confront what they define as problems in program policy or implementation (Kleiber and Light 1978). In general, the sources and accuracy of data on the economic component of structural context may become the focus of controversy which continues into the evaluation process and beyond (Levitan and Hurwitz 1977).

Research on the ORGANIZATIONAL COMPONENT of structural context entails critical examination of many of the "givens" of program operation: a) selected historical background on the evolution of the program, b) aspects of the present organizational structure of the program (size, hierarchical arrangement, service delivery system), and c) sociological documentation of the general management and negotiation models followed by key policymakers and program managers. Ordinarily policy researchers have little difficulty in locating historical and descriptive data on program organization through documentary sources, accessible files and cheerful interviews with program managers and staff members. It takes much greater reliance on observation, access to management files on internal policy and decision-making, and in-depth interviews with policymakers and managers to establish their general management strategies and negotiation models (Dalton 1959). However, without critical analysis of this component of structural context, policy researchers may implicitly accept an experimentalist model of evaluation research which defines organizational process as essentially irrelevant to the study of program impact. Systematic research on the organizational component of program development and implementation overcomes the tendency to treat the service delivery process as a "black box" that intervenes between program input and program output.

The main reason for examining these components of structural context is to develop a framework which assists the policy researcher in linking specific aspects of structural context with the arrangements for program evaluation research negotiated at some point in the operation of an ongoing program, or
at the conclusion of a program with a limited existence. In the following section I present in some detail the issues around which those program evaluation negotiations most frequently take place.

Analyzing Basic Areas of Negotiation

Following Strauss (1978:6) we may consider the structural context of a program as the "background" and the negotiations entailed in the production and use of a program evaluation study as the "foreground" problem in this analytical framework. In approaching the concept of negotiation context Strauss argues that there are "many specific kinds of negotiation contexts pertaining to interaction among negotiating parties" (1978:99). His typology is related to permutations of the properties discussed previously in my overview of Strauss (section 3.2). In formulating the framework for analysis used here I continue to rely on the concept of negotiation context, but I would suggest that there are identifiable problems in the production and use of program evaluations which can be interpreted as a series of negotiation contexts. These are the six issues in the theory and practice of program evaluation introduced briefly in Chapter 1 (section 1.2). I propose that each issue become the focus for description and analysis as a negotiation context, so that in any case study the program evaluation process may be "taken apart" systematically and negotiation contexts examined in terms of the properties outlined by Strauss (number/experience/representation of negotiators, timing of negotiations, balance of power, stakes, visibility, number/complexity of issues, legitimacy boundaries of issues, and alternative modes of action).

Working from the theoretical and empirical sources outlined earlier in this section, I want to present some of the contemporary research on the range of negotiations which may go on in the process of planning, producing and using a program evaluation document. Each of the following basic areas of
negotiation, or negotiation contexts, becomes the focus for developing a set of working arrangements during the course of the program evaluation process: (1) delineation of major actors, (2) organizational placement of program evaluation work, (3) general research strategy, (4) research model and methodology, (5) fact-making, drafts and the final research report, and (6) planning for research utilization.

(1) **Delineation of Major Actors**

Program evaluators operate most often in what I have defined as a three level work environment. This involves: a) a funding agency/department which is executively or legislatively required to contract for independent evaluations of programs under its jurisdiction; b) a work base, usually a profit-making firm or non-profit research institute or academic institution which employs program evaluators on a full-time, part-time or consulting basis; and c) a research setting which may be an entire agency, various site locations of a large-scale program or a specific project which the program evaluator has been hired to assess. A substantial amount of control and communication takes place which focuses on the funder/evaluator relationship. The formal guidelines for this relationship are generally set out in the contractual arrangements specified for the evaluation process. The contracting process defines formally the main actors whose interests are to be taken into account in the production and use of the evaluation research results. It is important to understand the circumstances in which the ideological position taken by the management of the funding agency, and the procedural/timing details of the federal procurement cycle can become major obstacles to the "representativeness" of research interests and research results (Fetterman 1981).

There are problems in the work base which may also affect the production and interpretation of research results: reporting demands from the funding
agency, the organizational arrangement and division of labour within the workbase itself, and of course the difficulties of conflict, stress and fatigue which can develop while evaluators are in the research setting. These problems are further complicated when the evaluator(s) must stay involved for some length of time in a research setting which is a large scale program with a relatively extensive policy mandate, a well developed administrative bureaucracy, an elaborate system of service delivery and a large population of clients/users/consumers. It is in these settings, e.g. school systems, health centers, employment preparation programs, that many of the decisions made in the controlled environment of the work base must be reassessed by researchers in the field and renegotiated with powerful members of the program being evaluated.

In this outline of the major actors represented in the work environment of program evaluators there are two other groups who must be mentioned here, the media and the public. They may play a very important role in some evaluation studies, while in others their role may be relatively minor. In policy studies which are designed to be internal or confidential, the media and the public generally have no input unless information is "leaked" and the program is a controversial one. Clearly, the significance of media and public participation in aspects of the policy process varies enormously depending upon the scope, political sensitivity, and accessibility of information about a program (Cronbach 1980:121). Analysis of reform programs from the last two decades suggests that public definitions of social needs, policy approaches, and evaluations of government sponsored programs are being increasingly moulded by media exposure and manipulation. Many provincial and federal Ministries hire advertising firms to "sell social problems" (Stern: personal communication). This process of media "hype" may further enhance existing tendencies among policymakers and program managers to support social issues that are politically "safe" and non-controversial. Alternatively, media campaigns may sensitize both policymakers and the public to
policy research which has previously received little attention (e.g. violence against women and children).

(2) **Organizational Placement of the Evaluation**

Program evaluation is a relatively recent management tool for setting policy in government programs. Originally it was taken from the private sector of the economy and the transfer occurred in the United States in the post-McNamara years (Deutscher 1976a). Policymakers trained in a private sector saw the growth and development of program evaluation as being linked with the central concepts of management/organization theory. Program managers were advised to examine the evaluation function within the context of coordinated management of complex organizations. From a perspective within organizational theory, the management of complex organizations could be described in terms of a four stage cycle: planning, decision-making, implementing and evaluating (Crowfoot 1977). The most common forms of program evaluation involved program audits and/or program reviews done by knowledge production units within funding agencies or programs themselves.

However, over time, the process of evaluation in large-scale government funded programs has tended to shift to a structural locus outside that original cycle. Evaluation, as a basic component of organizational development and change, has been moved some distance from the day-to-day operations and control of the management core of various large-scale public programs. The rationale frequently given for this structural shift was that outside evaluations would make programs more efficient, more accountable to the taxpayers, etc. Furthermore, the influence of policy science has encouraged the relocation of program evaluation research to organizations (firms, research institutes) whose expertise in the methodology and techniques of policy research is claimed to be superior to that of "in-house" researchers (Etzioni 1978:4). Thus, organizational
placement of program evaluations has become a controversial issue, frequently expressed as a dispute between the relative merits and demerits of having evaluation, and the evaluator, "inside" or "outside" the program being evaluated.

In some of the literature "internal" and "external" are treated as if they were dichotomous. The external evaluator is said to have greater "objectivity", frequently defined in terms of credibility, which is thought to derive from a lack of personal commitment to aspects of the program under review. The objectivity of the evaluator(s) is reified in the report which is produced and goes into "document time" (Smith 1974b). Generally the major strength of the internal evaluator appears to be in her knowledge of the program, its history and specific circumstances. This information is considered necessary for a balanced, in depth evaluation of a program. Whereas the objectivity of external evaluation is equated with distance, internal evaluation equates objectivity with depth of knowledge or comprehension of "the nature of reality" (Franklin and Thrasher 1976:132). This treatment of internal and external evaluations as being somehow distinct and virtually contradictory (in terms of location, process and product), assumes that programs are complete, homogeneous entities, where the evaluator is either inside or outside relatively impermeable organizational boundaries. However, it has been suggested that if we look at programs as a series of systems levels, this permits us to see program evaluations as lying along an internal-external continuum (Franklin and Thrasher 1976:130). Thus, "if the evaluator, or the ownership of evaluation, is seen as a component of that system level being evaluated, the evaluation is considered internal. If, however, the evaluation or evaluator is seen as a representative of a suprasystem, then the evaluation is seen as external" (Franklin and Thrasher 1976). The relative location of a program evaluation along the internal-external continuum may be related to such factors as the management ideology and legislative mandate of a particular program. In addition, organizational placement may enhance or limit the effective utilization of research results and policy recommendations.
The General Research Strategy

The "value screens" which orient policy analysts to different causes and means of resolving contemporary problems (Rein 1978:80), also orient the research and interpretation activities of program evaluators and other major actors in the policy process. However, program evaluators who come out of the policy science tradition tend to downplay or dismiss the influence of their own ideological perspectives (Boruch and Riecken 1975:10). Furthermore, upper level bureaucrats in funding agencies and managers of large-scale programs tend to deny that shared theories of intervention affect problem setting for the program evaluator, shape the meaning of the findings obtained, and determine the implications of such findings for the policy process (Fetterman 1981). Where all the major actors in the evaluation process do share a policy perspective and theory of intervention, negotiations often center around details of methodology and presentation of research results. It is where competing frameworks, differing theories regarding social problems and social change, are strongly represented in the evaluation process that questions regarding the evaluator's general research strategies become the focus of controversy (Siegel and Doty 1978).

Problems regarding general research strategy are compounded in the evaluation of large scale, multi-site programs. Often evaluators hired by the funding agency prefer to follow a consensual strategy, obtaining only the data requested and emphasizing those results which suggest minor program modifications to enhance overall program efficiency. However, differences in client/participants/consumers, along with variations in program management, staffing, and delivery may encourage some elements of the contentious approach among field researchers who wish to act as moral witnesses or social critics in their evaluations of these programs (Datta 1976:130). Finally, careful tracking of the long term policy shifts of a reform organization, and re-analysis of program evaluations done at different phases, may be part of an evaluator's paradigm-challenging
strategy designed to assess fundamental changes in the balance of power and access to resources among major actors in the policy process (Hessler and Beavert 1982).

(4) Research Models and Methodologies

Specific programs are intended to translate a broadly defined theory of intervention into concrete actions. In the development of these programs there are major differences regarding that translation process—from theory to implementation to evaluation. One position maintains that in order to produce better information about which government interventions work and which do not, it is necessary to design social programs deliberately as experiments (Boruch and Riecken 1975). Random selection becomes the crucial strategy and reliable, valid information about social processes and programs can be gathered through the experimental approach. For program evaluators that means an emphasis on immediate, discrete, quantitative measures of change, the strict requirement of pretest-and post-test data on program participants, and a concentrated effort to measure program impact rather than the dynamics of program delivery (Røss, Freeman and Wright 1979).

There is an alternative framework which proceeds from the assumption that "experimental evaluations on their own miss certain processes and outcomes crucial to an understanding of any program intervention" (Rein 1976:133). The main implication of this approach is that social intervention programs are seldom designed as experiments, certainly do not operate in the field as experiments, and therefore require a study design that includes prepared and sensitive observation of program process. The emphasis on observation (including in-depth interviews, recording of external forces that impinge on programs, etc.) as well as research design has developed as a response to three problems encountered in traditional, experimentally oriented studies (1976:134):
1. A tendency to concentrate on discrete, measurable changes, even if outcomes cannot be measured that way and what is measured may not be related directly to the purposes of the intervention;

2. The preference given to short-term changes, even when such a procedure imposes alien goals on a program and misses out on the long-term, slow-maturing qualitative changes that are more difficult to measure;

3. The tendency to evaluate programs in isolation from the context in which they must operate, ignoring characteristics of the school system or health care delivery system or labour market which necessarily impinge on the short and long term effects of any intervention program.

Policymakers, members at executive levels in funding agencies or government departments, and program managers may or may not express explicit preferences regarding the appropriate research design and methods for evaluating a particular program. But frequently aspects of the preferred research model are specified in the "request for proposal" which precedes the granting of an evaluation contract for large-scale programs. Again it is not a question of dichotomous alternatives, a completely positivist, experimental approach vs. an entirely interpretive, observation oriented approach. Rather, the evaluators and other major actors try to reach some agreement as to the balance of quantitative and qualitative data which will produce a set of usable findings and recommendations. However, there are many examples where extensive use of a traditional, experimentally oriented approach has essentially re-cast the purposes of a program through the process of evaluation. The best known of these examples is the Head Start program. The designers of Head Start saw the purpose of the program as the promotion of social competence, but in the national evaluation of the program there was an emphasis on cognitive development of the young program participants. This shift came about because the academic experimental psychologists imposed the narrower cognitive focus. In part this cognitive focus reflected their own interests, but also reading, vocabulary, and comprehension skills are much easier to measure than the fuzzier concept of social competence. It then
becomes evident the extent to which the values of the experimentalists, who see themselves and their research approach as being essentially passive, technical, and value-free, intrude at every stage in the process of research design, methodology and interpretation of findings. Questioning of the experimental model has become more widespread among policy researchers in recent years. An increased awareness of the significance of research models and methodologies in the shaping of research findings and interpretations seems to be making its way into the recent literature on program evaluation (Cronbach 1980).

(5) Fact-Making, Drafts and the Final Research Report

Researchers who are in the process of constructing a document called a "program evaluation" are simultaneously engaged in fact-making at three interrelated levels. At the micro level there is the process of telling (verbal or written) about some event or activity that has been classified as observable and relevant by the categories and conceptual procedures employed in the research design. These "data" which are found (e.g. in existing files on clients) or created (e.g. through questions about a respondent's work history) are then read as a series of facts (Smith 1974a). At an intermediate level of abstraction the researchers use a limited number of accepted methods for assembling these facts into groupings that are described with varying degrees of detail and summarized (statistically or in paragraph form, or both) as the findings of the research. Finally, at a macro level, following procedures influenced by the contractual requirements of the sponsoring agency and the research strategy followed in the overall evaluation process, these findings are given global interpretation as a series of policy recommendations in the evaluation document. One final phase occurs when the document has been accomplished (produced), i.e., the process and producers tend to disappear (Smith 1974b). This is especially true for program evaluation documents which present findings and recommendations inside a cover with a corporate logo.
There are two fundamental issues which affect all levels of the fact-making process:

1. The practices and limits of quantification, and its implications in the construction of findings and formulation of policy recommendations;

2. The context of production and reading of the factual documents required by most program evaluation contracts.

The historical events and ideologies of an era are reflected in particular settings where "data" are generated. In any specific instance it is what happens in the time interval between the event and its final recording that gives us a basic guide to the various fact-making practices which are being employed. In the conventional positivist paradigm of evaluation research which relies on quantitative analysis, this critical time interval and the operations which go on within it go virtually unrecognized. For example, systematic critical recognition of the limited reliability of official statistics (Cicourel 1964; Douglas 1976; Wheeler 1969) is ordinarily translated by positivists as a warning about "poor counting procedures" and the problems of human intervention in the data collection process. In addition, social scientists with heavy intellectual investments in high technology data processing tend to have limited concern with the preliminary manipulations which generated their input data.

However, in program evaluation and social policy research there is an alternative position which argues that the processes of counting, quantifying, rate-production and measurement cannot be taken as "givens" of the research process. Bogdan and Ksander (1980) present a comprehensive, systematic approach to the study of rate producing behaviour in terms of the concept "enumerology". Enumerology is the study of social processes by which numbers are generated and the effect of these processes on behaviour and thought (1980:302). From this perspective we are asked to approach statistics as something that "constructs reality", and directed to examine how counting actually occurs, how organizations function and the place of counting in everyday life. The significance of
enumology for the construction of program evaluations lies in the forceful directives it gives that we recognize the pervasive nature of quantification, while arguing that there is no need to eliminate quantitatively analysis from the policy research process. The authors suggest that we "balance our view of numbers as a method of objectively reporting reality by seeing them as part of the process of constructing reality" (1980:307). As program evaluators plan and write the drafts of the evaluation document their approach to the question of quantification is frequently reflected in the measures of program success/failure that are reproduced and used in the text. Identification of the processes by which rates were produced, clients counted, and figures generated provides document readers with additional evidence on which to accept or reject the findings and interpretations presented in any policy research document.

The positivist perspective on policy research documents such as program evaluations suggests that we view the product as an objective, independent artifact which can be used by anyone who has an interest in the topic under review (Etzioni 1978:4). A sociology of knowledge perspective argues that factual documents are social constructs which have a fundamental relationship to the practices of governing, managing and administration in contemporary society (Smith 1973a, b). Furthermore, most factual documents are not made to be detachable from specific organizational contexts of interpretation. It is this problem of the degree of "detachability" which so often becomes central in the transfer of evaluation documents from the producers to a variety of other contexts of interpretation (e.g. policymakers, the media, the public, and of course, the funding agency).

One of the main strengths of a relatively traditional experimental/positivist approach to program evaluation research is the apparent detachability of the product. There tend to be fewer confrontations among the main actors where the positivist model for the accomplishing and reporting of research results has
been the dominant paradigm throughout the evaluation process. However, in work situations where there is an interdisciplinary study team, the program evaluation document may go through several internal drafts in order to reconcile varying research approaches. Then the authors must negotiate that document through critical reviews by policymakers, the funding agency, program managers and participants, and on occasion the media and the public. Usually, despite controversy over drafts, a final stabilized version of the text is produced. This filing of the text in official form, whether by publication or by some reproducing process internal to the organization, has the effect of making available a version held to be the same on every occasion for all readers. Contemporary bureaucratic and professional practice is dependent on the continual production of a series of accounts which are entered into document time.

(6) Planning for Research Utilization

In approaching the problem of the potential use of policy research studies, program evaluators may not assume that funding agencies and policymakers will automatically process the information made available in the evaluation document. With regard to agency representatives who fund large-scale evaluations there is substantial evidence (Cronbach 1980:232) that much of their decision making activities are designed to be self/agency protective. They are frequently concerned with information that reduces their own vulnerability and helps to maintain the organizational status quo for evaluation sponsors and for programs. Therefore, program evaluations which proceed from the contentious, or paradigm-challenging approaches, produce findings which are "counter-intuitive", and support program activities which have been previously criticized, are highly likely to end up on the shelves of a storage closet in the funding agency office (Saroff and Levitan 1969).
There is little systematic empirical research on the use of social science knowledge in policy formation at provincial (state) or federal levels of government. One of the most widely quoted empirical studies of policymakers and their decision-making orientations has some direct implications for program evaluators engaged in establishing strategies for assuring that policymakers actually use their evaluation document. The research by Nathan Caplan (1977:183) emphasized that "knowledge utilization of any kind does not occur in a vacuum". Therefore his purpose is to derive (from interviews with 204 decision-makers) a minimal set of conditions necessary for the utilization of social science knowledge in policy formulation at upper levels of government power and responsibility in the United States. At the risk of oversimplifying the results, Caplan indicates that utilization is most likely to occur when (1977:183):

1. The decision-making orientation of the policymaker is characterized by a reasoned appreciation of the scientific and extra-scientific aspects of the policy issue.

2. The ethical-scientific values of the policymaker carry with them a conscious sense of social direction and responsibility.

3. The policy issue is well defined and of such a nature that a "best" solution requires research knowledge.

4. The research findings share the following characteristics: (a) they are not counterintuitive; (b) they are believable on grounds of objectivity; and (c) their action implications are politically feasible.

5. The policymakers and knowledge producers are linked by information specialists capable of coupling scientific inputs to policy goals and objectives.

Caplan's findings focus on specific characteristics of the policymakers (decision-making orientation, ethical-scientific values), the policy issue, characteristics of the research findings, and the relationship between policymakers and knowledge producers. These categories of analysis suggest a kind of "checklist" for program evaluators who try to anticipate the possible uses of their research document and develop strategies to enhance this acceptance process. In particular Caplan's analysis of the decision-making orientations
of the policymakers he interviewed puts the issue of the use of social science knowledge in the context of the user's own orientations to knowledge. The basic premise of that part of Caplan's study was that "the ways in which policymakers process information appear to have different consequences in determining the amount and kinds of knowledge used in arriving at a policy decision even after variables such as rank and department are statistically controlled" (1977:186). Respondents could be classified in one of three information-processing styles—clinical, academic or advocacy—on the basis of their own descriptions of how they used knowledge pertaining to the scientific and extrascientific aspects of the policy issue under consideration. Caplan uses "scientific" to refer to matters bearing on the INTERNAL LOGIC of the policy issue (gathering, processing, and analysis of the most objective information for reaching an unbiased diagnosis of the problem) and "extrascientific" as bearing on the EXTERNAL LOGIC of the policy issue (political, value-based, ideological, administrative and economic considerations).

The clinical orientation, favoured by policymakers with medical backgrounds, made greatest use of empirically based information to deal with the internal logic of the problem, then gathered information regarding political and social ramifications of the policy issue to deal with the external logic of the problem. To reach a policy decision, they weighed and reconciled the conflicting findings/recommendations of both types of information.

The academic orientation, favoured by those with Ph.D.'s who are often experts in their fields reflected a preference for information devoted to the internal logic of the policy issue. These policymakers were much less willing to cope with the external realities which confound policy problems, frequently considering such problems as a menace to the prestige and standing of their expertise. Consequently, they use scientific information in moderate amounts and in routine ways to formulate and evaluate policies largely on the basis of scientifically derived information.
The advocacy orientation was preferred by federal officials with legal backgrounds who were much at home in the world of social, political, and economic realities. Their use of social science information was limited, and use was dictated by extrascientific forces to the extent that members would intentionally ignore valid information that did not fit the prevailing political climate. These policymakers were preoccupied with the external logic of a policy issue, and when used in that context, scientific knowledge served to rationalize a decision made on extrascientific grounds.

Caplan's delineation of these styles of information processing suggests that some of the difficulties experienced by academically oriented knowledge producers in their encounters with policymakers may derive from significant differences in information-processing style among the two groups. Furthermore, in examining the ethical-scientific values of policymakers Caplan found that:

"... what seemed especially crucial in the decision-making process was the application of a value-laden appraisal of the possible consequences of the policy decisions. This is not to deny that many respondents provided citations to specific social science information, particularly research, and emphasized its importance to the decision-making process. But such 'hard' knowledge (research based, quantitative, and couched in scientific language) was usually only of some instrumental importance, and the final decision—whether or not to proceed with a particular policy—was more likely to depend upon an appraisal of 'soft' knowledge (nonresearch based, qualitative and couched in lay language) considerations of the possible social consequences of the policy decisions" (1977:188).

One may question the hard/soft dichotomy; however this finding does fit generally with Rein's discussion of the significance of "value-screens" in providing a framework for the interpretation of policy issues and policy research. Not only does the contemporary policy science approach to knowledge production neglect the value context of policy research, but, as Caplan points out, present theories of knowledge utilization are often limited by their overreliance on an assumed pattern of knowledge use involving hard information (scientifically produced). One implication for program evaluators is that a strategy to increase
utilization of evaluation results would likely involve a clearer statement of the evaluators' value position and increased reliance on qualitative information presented in non-technical language.

The Caplan study concluded with an effort to test out the significance of various commonly held theoretical positions on the use and nonuse of scientific knowledge. In analyzing the attitudes of the respondents Caplan found that theories of underutilization with the greatest degree of explanatory power were those which emphasized the existence of a gap between social scientists and policymakers due to differences in values, language, reward systems, and social and professional affiliations (1977:194). These findings strongly suggest that "social scientists would be well advised to pay particularly close attention to the utilization theories that stress the lack of interaction among social scientists and policymakers as a major reason for nonuse" (1977:195).

This emphasis on the need for interaction, and in particular the influence of the personal element has emerged in several other studies of the policy process. Cronbach has stressed that influence comes from engagement, not detachment (1980:153). And a number of research reports in the Weiss collection (1977) on the use of social research in public policymaking stress that what really counts is the interest of one or more officials in learning from the research and, simultaneously, the effort put in by the researcher to gain attention for what she knows. Patton and others (1977:153) summarize this position regarding the dynamics of knowledge utilization:

"The specifics vary from study to study but the pattern is markedly clear: Where the personal factor emerges, where some person takes direct personal responsibility for getting the information to the right people, evaluations have an impact. Where the personal factor is absent, there is a marked absence of impact. Utilization is not simply determined by some configuration of abstract factors; it is determined in large part by real, live, caring human beings."

The limited amount of empirical work specifically concerned with utilization of social research in the policy process documents the complex set of factors
which may enhance or prevent such use. The work of Caplan and others suggests that program evaluators must consider specific characteristics of the policymakers, the policy issue, characteristics of their research findings and their personal relationships with the policymakers in planning strategies for assuring a positive reception is given to their evaluation report.

**Implications for Case Study Analysis and Model Development**

As indicated previously (section 3.3), the framework presented here has been designed to both extend and break down the program evaluation process as it is accomplished in a variety of field settings. In the course of examining the four case studies discussed in Chapters 4 and 5, I will be describing and analyzing only those components of structural context and basic areas of negotiation which emerged empirically as the foci of extensive discussion and interaction among the negotiating parties. For example, in the first two case studies the organizational placement of the evaluation work was not subject to negotiation, although the hiring process in each case was significant and offered a contrast between overt and covert forms of interaction. Also, in comparing the pairs of studies, one becomes aware that the question of research utilization was essentially a non-negotiable issue in the first set of studies, while in the second set there were somewhat similar obstacles to utilization that were partially overcome.

The detailed documentation of both program background (components of structural context) and the program evaluation process (basic areas of negotiation) is intended to provide the "thick description" essential to build up an expanded set of models for doing policy research with relevance to policy intervention objectives and program context. Therefore, in analyzing the overall pattern of working arrangements which emerged in each case study setting I will be emphasizing what appear to me as patterns of evaluative knowledge production and use
which are linked with components of the structural context of each program. In addition, I will be looking at shifts in the structural context of a specific policy area and examining their impact on the pattern of evaluative knowledge production and use which emerged during that period of social change. Following Strauss' (1978:6) background-foreground metaphor, I want to use the proposed analytical framework to help in formulating an expanded set of models for policy research which are essentially stylized reproductions of the policy research process that retain the strongest elements of the policy intervention "picture".
CHAPTER 4: CHANGING STRUCTURAL CONTEXT AND PATTERNS OF COOPERATION AND ANTAGONISM

As research director of two evaluation programs concerned with the rehabilitation of housing and neighbourhood amenities (undertaken 1975-1977), I worked with two different research teams to observe, document, analyze and evaluate limited aspects of the change in Canadian housing policy during the decade 1968-1978.

The evaluation documents (Levitan and Hurwitz 1977; CPAC (B.C.) and Ptarmigan Planning 1977) focussed on the social settings in which the programs were implemented, details of the goals and day-to-day operations of the programs, and analyses of the social and economic consequences of the programs. In this chapter, however, I look at the program evaluation process from a different perspective from that of a research director of program evaluator; using the conceptual framework developed in Chapter 3, I describe and analyze the negotiation processes through which those evaluation documents were commissioned and produced, and the "policy limbo" into which they were, unfortunately, delivered. This new perspective involves: (1) an examination of the structural context in which the programs emerged, i.e., the policy-making, planning and implementation activities of those key actors concerned with housing as a social problem; (2) a description and analysis of the negotiation context in which the work of producing the evaluation document was accomplished; and (3) a discussion of each case study in terms of the overall pattern of knowledge production and use which characterized the program evaluation process.

Given this analytical agenda, I begin by outlining selected historical/political aspects of Canadian housing policy since this provides background on the structural context for both studies. An examination of housing policy will also introduce three of the most significant negotiators in each case
study: the federal government (through Central Mortgage and Housing Corporation), the provincial government (through the Ministry of Municipal Affairs) and the municipal government (through Departments of Planning and Social Planning). In addition, we become aware, almost immediately, of the complexity of the issues involved in the development of housing policy and the ongoing negotiations concerning legitimate terrains and boundaries which have been carried on continuously since World War II.

4.1 The Structural Context: The Policy Shift from "Slum Clearance" to "Neighbourhood and Housing Rehabilitation"

Spheres of Control

Canadian housing policy, though often referred to in ways that reify it as some "entity", has always been a shifting series of pieces of legislation and sets of regulations administered through negotiations among representatives of the three levels of government. From an historical perspective, government participation in house financing began in 1918, when the federal government made $25 million available to the provinces under the War Measures Act for re-lending to municipalities for housing purposes (Neufeld 1972). The Dominion Housing Act (DHA) of 1935 gave the federal government a continuing role in residential financing, with the objectives of improving housing standards, creating jobs in residential construction and expanding access to mortgage credit (Mitchell 1979).

CMHC economist E.C. Mitchell (1979), currently Coordinator of Economic Analysis and Planning at the British Columbia Regional Office of the Central Mortgage and Housing Corporation (CMHC), dates the first step of federal "intrusion" into the provinces' sphere of control over the housing situation to the National Housing Act (NHA) of 1938. Specifically, under this Act, the federal government could make loans directly to local housing authorities
or private limited dividend organizations, bypassing provincial and municipal authorities. Then a 1944 amendment to the NHA brought increased access to mortgage credit along with the provision of aid to municipalities for slum clearance. Again, Mitchell interprets this as another intrusion into the "provincial sphere". From his perspective, the parliamentary establishment of the CMHC in 1945 was intended not only to administer the NHA but also to "put some room between the federal government and the NHA administration and as such would pacify any provincial concerns". His language illustrates how the basic issues regarding "spheres of control" were perceived by the federal level. The amended Act of 1945 permitted CMHC to make direct loans to individual families for housing, an important change which eventually lead to extensive federal involvement in housing production. In 1949, further changes to the Act brought the first joint federal/provincial ventures into agreement for the provision of public housing through joint funding (75:25 per cent basis) and later through a mortgage insurance arrangement in 1954. At that point federal intrusion was complete.

Policy Issues: Production vs. Distribution of Housing

In addition to the emerging pattern of federal/provincial conflict over housing policy, other patterns of conflict became evident, particularly over the issue of quantity/quality. During the post-war period and into the 1960's, "the primary thrust of federal involvement in housing was production oriented, and this production orientation focussed on providing housing for moderate and middle income Canadians" (Mitchell 1979). From the CMHC perspective, this was the primary concern of society at the time and social housing issues (low income, handicapped, elderly, native) were addressed only on a limited basis. It was only after the needs of "middle Canada" had been satisfied that attention and funding were made available for social housing
Several groups took issue with the production-oriented, middle-Canada approach to housing policy. One, the Ontario Association of Housing Authorities (1964), agreed that housing performance under the NHA had been production-oriented. But from their perspective the preferred policy approach would make government funded housing distribution-oriented, i.e., designed to serve the housing needs of those groups in Canadian society whose economic position placed them outside traditional housing markets. They accused government of developing "a quantitative operation qualitatively devoid of broad social objectives and economically inaccessible to many Canadians. The production of new houses should be a means to an end, not the prime policy objective." (p. 49).

To some extent the federal policy in the 1960's attempted to address the issues of social housing through increased involvement in urban renewal, slum clearance, and low rental housing projects. At the same time the provincial levels of governments became increasingly concerned with social housing issues. As Mitchell sees the problem developing:

The provinces, led by Quebec, were quite concerned about the intrusion of the federal government into what they regarded as their own sphere - social policy. In their view, the federal government was being used to get in through the back door in the provision of social assistance to people, and this is a provincial responsibility (1979:3).

But, apparently the federal government did not share this point of view. It continued with a variety of programs which were attracting increased opposition among citizens in all the larger urban areas across Canada. The Hellyer Task Force on Housing completed a cross-Canada tour in 1968 and the case against urban renewal with redevelopment had been conclusively made in each city where hearings were held. In early 1969, therefore, the federal government froze its funding of projects carried out under urban renewal. This appeared to indicate the beginning of a basic change in attitude toward the
problems of inner cities. Leaders of various provincial governments apparently reached similar conclusions about the high social and economic costs of urban renewal. Representatives at all three levels of government began to look for alternative strategies to deal with the housing problems of low income, inner city, and eventually rural and native, groups.¹

The policies (or lack of policies) of government in the area of housing came under the most intense and controversial scrutiny in the Dennis-Fish report delivered to CMHC in June 1972. Without disclosing the political rationale (only lack of a French translation) CMHC failed to release the 3,000 copies of the report it had funded. The study was then printed in book form by its authors, Michael Dennis and Susan Fish, (1972) as Programs in Search of a Policy: Low Income Housing in Canada. Many representatives of government, housing groups and the public showed interest in the Dennis-Fish Report. Their critique of Canadian housing policy can be summarized as follows:

Housing policy in Canada has been directed solely at starts. Its aim has been to increase the total stock of 'decent, safe and sanitary accommodation' to the point where there is sufficient adequate housing for all Canadians, demolishing substandard housing and replacing it wherever possible.

Little or no concern has been shown for: the distribution of either the newly produced or existing stock; the price of that stock and the ability of consumers, and of low income consumers in particular, to afford it; the environmental quality of new housing produced; the condition of the existing stock, except 'slum housing' which would have to be destroyed and replaced; the right to free and dignified use by the consumer of his home.

Instead reliance is placed on the market to allocate the stock, set the price, determine the level of quality, and protection of the position of the low income housing tenant is left to the provinces. The only minor shift which has occurred to date has been the recent expansion in the last three years of the public housing and low rental housing programs, and a lesser attempt at assisted homeownership. Within those programs the emphasis is very much on quantity rather than quality. Units produced under these programs constitute a minute portion of the total housing stock (some 2 per cent) and come nowhere near meeting the need. The vast majority of low income households are left to the vagaries of the market (1972:1).
The Evolution of Rehabilitation Policy

Between 1944 and the early 1960's, a major shift in policy regarding rehabilitation of existing housing stock occurred. Although the 1944 National Housing Act was "An Act to promote the construction of new housing, the repair and modernization of existing houses, and the improvement of housing and living conditions", the emphasis on new construction eclipsed the equal rating given to the repair of existing houses. The federal government in general, and CMHC in particular, showed little concern for the condition of existing housing stock. In 1956, for example, the minister responsible for housing recommended that new housing projects should be directly associated with redevelopment; that is, the demolition of existing units to be replaced with new construction. All authorities were much more interested in increments in housing stock reflecting a positive growth approach and a confidence that continued growth would be the solution to Canada's housing problem (Dennis and Fish 1972:296). To them, the urban renewal provision of the NHA assumed that urban renewal simply meant clearance. As our later discussion of the Strathcona Rehabilitation Project will illustrate, this definition of "urban renewal" persisted well into the 1970's. Prior to 1964 the NHA authorized the minister to enter into an agreement with a municipality to provide a grant to help pay the cost of acquiring and clearing an area of land. That agreement had to state that the municipality would acquire and clear the area and then sell it either to a limited dividend housing company or to the federal/provincial partnership for the construction of a public housing project.

Finally, in a 1964 amendment to the act, the Corporation, CMHC, was permitted to pay one half of the cost of acquiring and clearing or improving lands and buildings in the urban renewal area. This provision, however, only permitted assistance for the improvement of buildings in an urban renewal area,
and the assistance could only be paid where the building was acquired and improved by the province or municipality. Dennis and Fish note that this effectively prohibited the Corporation from making grants or loans to private individuals, even in urban renewal areas, for the improvement of their homes (1972:294). As well, the amendment permitted CMHC to insure loans made to the owners of housing projects in urban renewal areas if: (a) the housing project met the requirements (or when repaired or improved would meet the requirements) of an urban renewal scheme for the area acceptable to the Corporation, and (b) the housing project met the Corporation's housing standards. Problems with that subsection arose because it was uncertain whether it applied only to the purchase or purchase and repair or repairs alone; because of these confusions it was repealed in 1967.

The emphasis, both at the federal and provincial levels, was always on the building of new housing and the demolition of housing considered as slums. As Dennis and Fish explain:

There is a popular misconception that most occupants of poor housing are tenants, frequently in large multiple buildings. In fact, this is not the case. Even with the spate of high rise construction in metropolitan centres in the 1960's, more than 70 per cent of all Canadian households lived in single detached or single attached units in 1970. It is not surprising then, to learn that the majority of all dwellings requiring rehabilitation are single family, owner-occupied units (1972: 300).

This misconception characterized the attitudes toward slum housing and the removal of what was felt to be unsanitary living conditions. That is, removing people from those situations and putting them into nice, new, clean public housing projects would somehow change the lives of the people who had lived in this housing. It did change their lives but, as we know now from the work of Herb Gans (1962), it certainly did not change them for the better.

Although the initial CMHC programs did not emphasize existing housing, as early as 1960 it was realized that various types of existing housing could be used to assist low and moderate income households to become home owners
Then, in the late 1960's and early 1970's, CMHC became increasingly concerned with rehabilitation. That concern came in response to pressures that resulted from both the programs and initiatives of other government departments and from the complaints of citizens' groups that became increasingly disaffected with the continued emphasis on redevelopment, whether public or private. There began a gradual movement toward a notion of rehabilitation of parts of the existing housing stock, instead of urban renewal (or what in the United States used to be called "people removal"). By the early 1960's the rehabilitation process had already been initiated and assessed as very effective in parts of the eastern United States and larger urban areas throughout the midwest and the west coast of the United States. It took fully ten years or more for government policy-makers at the federal level in Canada and at the CMHC to recognize that this might be a much more effective way to deal with the problems of older inner city and rural housing stock.

This represents a brief overview of the evolution of Canadian housing policy up to the point where the new rehabilitation policy was introduced, and the first of the programs used as case studies, the Strathcona Rehabilitation Project, was funded by the three levels of government. This project and the NIP/RRAP programs reflect successive stages in the changing structural context of Canadian housing policy during the 1968-1978 decade.

4.2 Case: Stage II Evaluation of the Strathcona Rehabilitation Project

According to its original government representatives, the Strathcona Rehabilitation Project (SRP) was set up in 1971 to serve as a model for the implementation of a newly evolving urban planning policy. It was developed by the Strathcona Property Owners' and Tenants Association (SPOTA) in
cooperation with the three levels of government and emerged as the first of a series of eight pilot projects, located across Canada, designed to explore the possibilities of citizen participation in community rehabilitation (CMHC Task Force Working Paper 1971; N. Bean 1975). Government representatives considered the SRP a case study in policy planning and program implementation for a new approach to housing rehabilitation and neighbourhood improvement. The project experimented with a revised form of administration through a four-partner committee, the Strathcona Rehabilitation Committee (SRC), made up of representatives from three levels of government (CMHC, Department of Municipal Affairs, Vancouver City Planning and/or Social Planning Departments) and the community (SPOTA). (For a full discussion of the Project see Levitan and Hurwitz 1977).

Several aspects of the Project setting and early decision-making arrangements had important consequences for the negotiation context of the Project evaluation. First, there was a negative history of redevelopment in Strathcona and a consequent bitterness between the Strathcona community and several City departments, including Planning and Engineering. In addition, there was an almost total lack of any precedents or guidelines for a community rehabilitation scheme with a citizens' group as a full partner. This caused a great deal of mistrust and ambiguity at the outset. These feelings were compounded by the fact that the Strathcona Working Committee (the SWC was the predecessor of the SRP) started all its deliberations during the urban renewal freeze of the late 1960's but before a new policy emerged. The original imbalance of power and resources in the SWC had very important consequences for the final expenditure pattern of the Project. Thus, although under the SRP contract SPOTA was allocated money to pay honoraria to their representatives and technical advisors (architects), much of the critical decision-making regarding the "big ticket items" (underground services including sewage and water lines)
in the SRP budget had been made by the Strathcona Working Committee (the SRC's 1969-1971 predecessor). That original Committee made its decisions without undertaking any cost/benefit analyses and without SPOTA being able to hire its own engineering experts.

In an attempt to overcome some of these inequities of resources and skill, the Chairman of the SWC, the Director of Social Planning for the City, promoted the use of consensus in decision-making. He felt that voting would enshrine the adversarial system which had so long been operative, especially between the City and SPOTA. Instead he hoped to create an atmosphere of some trust and the beginnings of a cooperative partnership between government and citizens. The effect of this consensus model of decision-making tended to encourage an impression of greater agreement among the four parties than actually existed in practice. It also masked the coalition work that was frequently done behind the scenes (among technical experts on government staffs) and permitted the City representatives to "drag out" negotiations regarding those issues over which it wished to retain control.

All the government representatives were highly experienced members of their departments, and, with the exception of the City representatives, these men could make policy and procedural decisions without referring back to their supervisors. It was felt that senior government staff would be necessary to get the project going and to work out understandings. The SPOTA representatives were much less experienced at negotiations of this kind. Given that this was a working class ethnic community (predominately Chinese, with some Italians and Ukranians), only the relatively young or relatively old members of SPOTA were available for daytime meetings, especially before honoraria were available. But these representatives had a direct and solid mandate from the block groups and rank-and-file SPOTA membership.
These background factors of setting and decision-making affected the course of the Project and the evaluation. The negotiations of the SRC went on for several years, from 1971 to the middle of 1977, and there are minutes from more than fifty meetings. And yet, the ideological framework and financial structure of the project was almost totally developed in the period which pre­ceded the SRC when "redevelopment" was strongly entrenched as the policy at Vancouver City Planning Department and the requirements of the Vancouver City Engineering Department took precedence over other possible financial commitments.

Background to the Stage II Evaluation

At some point early in the operational phase of the SRP it was agreed, among all the representatives on the Strathcona Rehabilitation Committee (SRC), that the project should be formally "evaluated". However, it was not determined just how that evaluation should be carried out. At first it was thought that CMHC Head Office would monitor and evaluate the SRP, as well as several of the other pilot projects across Canada. The Site Office staff began to collect data for monitoring as the project got underway. Later, CMHC funded SPOTA through a Part V Research Grant to carry out an evaluation from its perspective while the project was ongoing (SPOTA 1973, a,b). Subsequently, it was suggested that staff from the Vancouver Social Planning Department conduct an evaluation, but agreement could not be reached on this at the SRC. Finally, SPOTA requested that the SRC hire an outside, independent evaluator and suggested the research staff of the United Way of Vancouver. This suggestion was approved by the SRC.

A research design was devised which incorporated a two-phase study. Stage I was completed by Larry Bell of the United Way with the assistance of Richard Moore. The Stage II evaluation, done by Ptarmigan Planning Associates, constitutes the case study described and analyzed here.
Phase I: Pre-Contract Negotiations

The period leading up to the awarding of the Stage II Evaluation contract deserves attention because it established patterns that would persist throughout the entire program evaluation process. The large number of actors, the curious duality of formality/informality of interaction, and the protracted nature of negotiations -- all these characteristics emerged in the six months of preliminary discussion, proposal writing, and selection process which preceded actual work on the evaluation study.

The initial negotiations regarding the evaluation of the Strathcona Project began in April 1975 when I received a phone call from a friend who lived in Strathcona and was an active member of SPOTA on the SRC. She told me that there would be a proposal call for the Stage II evaluation of the SRP and recommended that we, Ptarmigan Planning, call Harry Pickstone, Deputy Director of Planning and the Chairman of the SRC. In response to our inquiry Pickstone wrote to Laurie Hurwitz, my partner at Ptarmigan (7 May 1975).

Dear Laurie:

I understand that you and Alberta Levitan have set up in the planning business and propose to take on the protective coloration of any neighbourhood in which you find yourself like your patron bird. I would be interested to receive a brief proposal from you for carrying out Stage II of the Strathcona Evaluation Study. I wrongly assumed, when I spoke to you over the telephone, that we already had developed an outline, but for your information I attach an outline of Stage I of Larry Bell's work.

Stage II will be to update Stage I and to attempt to measure the impact of the project on Strathcona as a community and on the individuals. The sorts of things which I mentioned over the telephone are:

1. Have the physical changes made the area more livable, e.g., less through traffic?
2. Has the project resulted in a stronger community spirit in the area?
3. What has happened to movement into and out of the area? Is it now more stable? Is it now proportionally more expensive than it was before?
4. Has the physical face lifting carried out with government assistance stimulated other physical improvements?

5. Alternatively, has the affect of the grant loan system engendered a feeling of paternalism?

6. What has this project done to neighbourhood facilities -- improved their quality, improved their use, or the reverse?

7. What has the project done to neighbourhood organizations - strengthened them or made them reliant on assistance?

The list is not by any means exhaustive but should serve to illustrate the sorts of concerns which should be addressed in the Part II Study. The purpose of the two evaluation studies is to serve as a basis of experience for designing other similar projects. It is appropriate, therefore, that the Part II Study should contain specific recommendations from the consultants which would be of use in setting up future projects.

I am soliciting proposals from various consultants which should include Vitae of the people who would be involved and some brief indication of how this study might be tackled. It is my intention then to discuss the matter personally with some of the people who have responded and I believe that the Strathcona Rehabilitation Committee will also wish to interview prospective consultants.

Pickstone closed by indicating that the next meeting of the SRC would be set up for May 28th and asked to let him know whether we were interested. The curious duality of formality/informality that was evident in this letter was to become one of the distinguishing characteristics of the negotiation process.

A second distinguishing characteristic of these negotiations was their protracted nature. We submitted a preliminary proposal for the Stage II Strathcona Evaluation on June 3rd 1975. The final decision to award the contract was not made until October of that year. The period in between was filled with numerous meetings and telephone calls, and an abundant correspondence. Copies of all letters went to CMHC, the City Planning Department, the SPOTA president, and to the SPOTA member who had originally contacted us. It became evident early in the process that everything had to be sent to the various key negotiators.

Requests for additional information accounted for much of the ongoing
negotiation process. In late June 1975, for example, the SRC decided that it should review Stage I of the study and then supply copies of the report to the potential consultants. This would provide them with an opportunity to modify or amplify their proposals. This was done on July 30th 1975. In October 1975 we submitted further detail about our proposed budget and time allocations for various tasks involved in the evaluation. We also stressed the importance we attached to the survey component of the proposal, provided some additional detail about the survey's content, and indicated that, in addition to furnishing ten copies of the final report to the SRC, we proposed to produce 300 copies of a three-page summary of the report for distribution within the community. (This was the first overt indication of our commitment to making the information available to community members who had participated in the program.)

The large numbers of actors involved in the negotiation process and the difficulties of assembling them to make decisions also explains the extended length of this pre-contract phase. Problems arose particularly in the summer months when, as A.D. Geach, the new Assistant Director of Planning and the SRC Chairman (he replaced Pickstone in June 1975) pointed out, it would be difficult to ensure that committee members would be able to complete their review of the Stage I Report within as short a time as might be possible during another period of the year.

In retrospect, one can see that, in many instances, the extended nature of this process worked to our advantage, by allowing us to fill in gaps in our research proposal and personnel. For example, in July 1975, in response to an enquiry about the economic analysis component of our proposal, we acquired the services of David Baxter, an economist experienced in urban housing economics and land use planning. In October, we added Gloria King, a Chinese cultural anthropologist, bilingual in English and Cantonese and experienced in working in Chinese communities in England and South Africa.
The decision about the contract was finally made on October 8th 1975. On that day, four planning consulting firms, including Ptarmigan, were interviewed. Each firm was allocated one hour in which the firm's representatives could present themselves and be "cross-examined". Laurie Hurwitz and myself, representing Ptarmigan, walked into the meeting to find representatives of the three levels of government, the formal representatives of SPOTA, plus an additional ten or twelve other residents of the Strathcona area who were very much involved in the rehabilitation program. We spent an hour being questioned by a total of more than 20 people. A substantial part of that time was spent discussing our backgrounds and work experience in Vancouver and elsewhere. Only then were we asked several questions about the general research approach planned and requested to elaborate on the various research sources we might need, especially files from the project site office and those at the CMHC Regional Office. At the end of that afternoon of consultant interviews we were told by a representative of SPOTA that we were to be given the contract. We never learned formally why we were selected, but our SPOTA informant indicated that our "style" and the makeup of our study team were preferred by all the representatives of that selection meeting. (For a full description of the study team see Appendix 1).

Also, at the end of that same afternoon, with little warning, we were immediately whisked off to a small banquet at a Chinatown restaurant. The other guests included government representatives and a substantial contingent of SPOTA members. This was the first of a long series of dinners that we attended during the Stage II evaluation. We knew we could not appear empty handed at this particular dinner, so we quickly rushed out and bought a bottle of rye, since we knew from other experience that this was the appropriate gift to bring to a banquet of this kind.
In a recent interview with my former partner we both remarked on what an unusual experience that entire day seemed, especially in retrospect. It was a rather amazing combination of settings and interactions: the relatively formal setting related to being interviewed for a contract, the curious experience of being interviewed by a large number of government representatives and community members in a process that combined "cross-examination" with political banter and ethnic wisecracks, the rather casual notification that we had received the contract, and the immediate acceptance and rapport we experienced at the dinner which followed. This particular set of experiences, which occurred at our first face-to-face encounter, was to typify the dual nature of the work of doing the evaluation. It was both very tough (aggravating, time consuming, frustrating) and very easy (good discussion, delicious meals and great rapport among members of the study team).

**Phase 2: The Evaluation Research**

We set out with great hopes. Work was to begin as early as possible, but not later than November 15th 1975, and was to be completed by May 16th 1976. The financial arrangements were pretty clear cut: the cost of our work was not to exceed $18,950; payment would be in three equal instalments (Geach to Ptarmigan, November 5th 1975). As well, the City Law Department had approved an addendum to our contract that allowed Ptarmigan Planning "the right to subsequently use the research or material in the final report either in part or in whole for the possible purposes of academic research and/or scholarly or professional publication or presentation. There would of course be appropriate citation as to the source of the material used" (Ptarmigan to Geach, November 21st 1975). Finally, we hoped for cordial relations with the other actors in the process. It seemed clear that our
best sources of information and the most cooperative element in the whole arrangement would be the SPOTA members and other key people in the community who were involved in making the program a success. But as well we had a sense that CMHC personnel would be as cooperative as possible, that the provincial authorities would allow us to use files and data. Only with the City Planning Department members, those who would be involved in the day-to-day detail, overseeing our activities, did we anticipate any friction.

By the end of December 1975, we had the study team, the main people together, and most of the administrative detail arranged. But things were taking a lot longer than anticipated; moreover this was the first evaluation this team had done together.

Our first progress report to the SRC was dated January 22nd 1976 and covered the period November 15th 1975 to January 22nd 1976. It detailed various kinds of things that we had done, and outlined what the anthropologist, the economist and the physical planner were doing. By that time we had added to our research team a University of British Columbia planning student as an intern and she was working in conjunction with us and the coordinator of the Strathcona Community Resources Board Social Needs Study to develop an interview schedule which would go to the social service agencies in the area. The letter also indicated our work schedule for the coming months, and in our best anticipatory style, we figured that we would have the first draft of the report done by May 15th 1976, not only the first draft, but the SRC comments and the drafting and presentation of the final report. On January 27th 1976 we received the first third of our payment. We photocopied it and had it put in a frame at one point!

In mid-February 1976 we decided to rent office space in the Strathcona area, knowing it would be easier to work right in the community. Since SPOTA was having some financial difficulties, we asked if we could rent
space from them. We moved in on March 1st for a period of three months with an option to rent for an additional three months if agreeable to all parties. It was in fact a good move, because we then were located where people could just drop in and talk to us. As well, there were several boxes of critical files at CMHC (which was moving its office at that point) and at SPOTA which might have been dumped in the trash had we not had the place to put them. These were files specifically on the program and background to the program.

As in the pre-contract period, we continued to make all other actors aware of our activities. For example, on February 19th 1976, we sent a letter to Geach listing the proposed topic areas that formed the basis for the questionnaire to be used with SRP participants. We asked him to contact us by February 26th if he had any comments or any specific topic areas and questions he would like included. We sent a similar request to J. Williams at the Department of Housing, J. Hines at CMHC, and T. Mesic, SPOTA president. Thus everybody was notified of the topics to be included in the interview schedule. All these transactions were highly visible and little covert work occurred in the negotiations themselves. We had a strong feeling that all activity could be up front, that nothing need be hidden. Only later did we begin to find that certain kinds of interviews had to be approached indirectly, that we had to use manipulation, a good bit of education, and other alternatives to straightforward negotiation in order to get our work done.

We completed our second progress report on April 7th 1976. We had made further additions to our team, hiring a research coordinator for the interviews, Jonathan Lau, a social worker originally from Hong Kong, who had worked in the Strathcona area. We had completed a whole batch of focussed interviews with SRP participants; the real estate analysis had been done and we had contributed to the design and implementation of a survey of Strathcona area social service agencies, directed by Michael Goldberg who
subsequently did the data analysis on the results of the interview schedules. Also we had completed an assessment of the SRP files on all the participating residents, and our physical planner had completed the research and evaluation of the engineering works.

But we were running behind schedule. A whole series of unanticipated tasks cropped up, forcing us to revise our time estimate and to detail the "factors involved in revised time estimate for completion of Stage II evaluation". These factors slowing us down included the task of producing the final interview questionnaire in both Chinese and English, a much lengthier process than we ever anticipated. As well, the length of individual interviews with key personnel at the SRP had been considerable, ranging up to about six hours in some cases. The original proposal had not envisioned undertaking a comprehensive survey of agency workers, or the extensive analysis of the files of all project participants, both of which were undertaken subsequently as part of a plan to increase our "data base". As it was, the survey of agency workers was not terribly useful, but the analysis of the files, particularly the financial analysis, did prove significant in terms of documenting the level of financial participation of people who participated in the SRP. At that point we had several tasks to complete: the survey of participants, the survey of non-participants and absentee landlords, some more focussed interviews, and additional parts of the real estate analysis to do, as well as the rest of the work entailed in assembling the report as a whole. However, we concluded our progress report with the optimistic statement that the evaluation would be "completed one month behind the projected schedule".

On June 4th 1976 we received the analysis and commentary on the economic implications of the project from our economic consultants. In the process of discussing that report we learned an enormous amount about alternative ways
in which the economic data could be analyzed, the problems of missing data
(real estate transactions uncollected at the City level), and the implications
of the financial statements summarizing SRP expenditures. (See Appendix 2
for an overview of all methods and data sources).

Five months later, on November 10th 1976, the draft report was finally
sent off. It had been written in discrete sections "to allow readers with
various levels of information about the project to focus on those areas of
particular interest to themselves" (Ptarmigan to Geach). We requested that
written comments from all four parties be sent to us at least one week prior
to the meeting on that report.

Difficulties in arranging that meeting were caused by the now familiar
problems of finding a mutually acceptable date for committee members to
consider the draft. As Geach wrote to Ptarmigan: "As a result, having
regard to the likely availability of members for a meeting in the New Year,
and present progress on review of the draft, it appears unlikely that it
would be feasible to have a meeting before the middle of February. February
16th appears at this time to be the most suitable date and I would request
therefore that members keep this date open for a meeting to deal with the
draft" (November 16th 1976).

Phase 3: The Evaluation Report

With the presentation of the draft report, the negotiation process
entered a new phase. Before that review meeting, called for February 23rd
1977, finally took place, we received several written responses to the draft.
CMHC did not produce a detailed written review, but both the provincial and
municipal authorities expressed their displeasure. The following comments
reflect the range and extent of that criticism.

The provincial representative on the SRC wrote to the SRC Chairman:
The report is very thorough and reflects deep research. However it is
verbose and so compartmented that several subjects are repeated many
times, e.g., the Community Centre, the underground servicing standard
and personal criticisms. Continual reference to 'bureaucrat' and
'bureaucracy' are tiresome -- there are other acceptable words.

By closer grouping of subjects, much of the repetition could be avoided,
thus making a more interesting and readable document. For instance,
the many dissertations on RM 3 servicing standards are made unnecessary
by the summation in Appendix 5, p. 139, where it may be overlooked.

Severe editing and consolidation are necessary to make an acceptable
report.

Some specific comments are attached. (Williams to Geach, December 15th
1976)

In a memo to the SRC Chairman dated December 17th 1976, the Assistant
City-Engineer, Departmental Services and Sewers, made these remarks regarding
the preliminary draft:

We have reviewed this draft and if its general tone really reflects the
feelings of the people in Strathcona, then the effort (and money) put
into SRC meetings, public meetings, and the site office have been
largely wasted. The 'SRC Government Representatives' are pictured
as working in bureaucratic isolation.-- the meetings we attended were
run almost as public meetings and decisions were by consensus. There
are repeated references to bureaucracy. The role of Council and its
Committees is ignored. We wonder how Ptarmigan's dictionary defines
bureaucracy.

There are 25 additional items in this memo, some of which have to do
with cost estimates which seem rather high, but most of which refer to the
significance and value of the underground services provided, and the extent
to which decisions relevant to these services had been made by the members
of Vancouver City Council and not the Engineering Department. Repeatedly
we were reminded of the role of Council and its Committees.9

In his memo to the SRC dated December 30th 1976, the Chairman, A.D. Geach,
offered criticisms of the length of the report, and some concern over the
accuracy of various figures on expenses. But the bulk of his critique
appeared on nearly eighty pages of the draft which had been photocopied and
specific changes in terminology, spelling and interpretation were presented
in great detail. Two sections of that memo deserve attention here:
(2) **Objectivity**

The tone of the report does not, in my view, reflect the objectivity which should be a requirement of a consultants' study of this kind. There are numerous instances of what can only be interpreted as a significant bias against staff of the various levels of government, in particular the City, who were working with new and untried concepts without the benefit of hindsight.

(3) **Terminology**

There is a fundamental error in terminology similar to the one which was identified in the draft of the Stage 1 report, and which, by agreement of the Committee, was corrected before the draft was finalized. This is the reference to 'urban renewal' instead of 'redevelopment' to describe the various projects in the area (Redevelopment Projects 1 & 2). The Strathcona Rehabilitation Project is an urban renewal project carried out under the Urban Renewal provisions of the National Housing Act that differs from the preceding activity in that it is an urban renewal project emphasizing conservation and rehabilitation rather than redevelopment. For consistency, this definition should be recognized and corrected, and I have annotated the draft in a number of places.

Obviously they were not pleased with the draft and made themselves quite clear on that. However no specific criticism was made of our major finding regarding the SRP budget; that is, that in looking at total costs, including the sum of owners' contributions, the federal government had covered 47.2%, the provinces 23.6%, the City 19.6%, and the homeowners 9.6%. In a program where approximately 60% of participants had total family incomes under $6000 (1972-1974), these low income participants had paid nearly 10% of program costs (including sewers that were designed to last 100 years).

SPOTA's response, while generally favourable, was also qualified. In a memorandum dated January 31st 1977, Jo-Anne Lee, executive director of SPOTA, offered the following general comments regarding the draft:

Members of SPOTA, active during the SWC and SRC meeting period, have discussed the report. Generally, favourable reactions were voiced. The report does include many of SPOTA's perceptions and feelings. However, we also found the report to be lengthy and repetitive. The main points need to be punched out more clearly, both in terms of the layout and style of writing. One gets lost in paragraphs which one is sure had been read before.
In retrospect, it is somewhat embarrassing to realize how little we understood about the process for writing and assembling a program evaluation document. The criticisms regarding length and repetitiveness were valid, and a basic "pruning" would have significantly improved its acceptability to the SRC members. But while we had been prepared for some questions regarding our basic methodology, this only emerged in the context of the problem of "length". We had, it seemed, been rather naive about the manner in which our interpretations of the data would be received.

After the initial shock, Laurie Hurwitz and I, as the principals of Ptarmigan Planning, mapped out a strategy to counter some of their comments so that we could emerge from the review meeting with some definite and clear agreements regarding the structure and content of the final draft. We began the process by sending a letter on February 16th 1977, before the meeting, in which we thanked the members of the SRC and the other concerned officials for the time they had put in reviewing and analyzing the final draft. In it we comment that: "The length of that report seems to be a matter of particular concern for all readers. We agree that for greater impact the final report should be substantially condensed, without losing valuable information. We are preparing a proposal on how we would suggest that the report be shortened and will bring this to the meeting on February 23rd. Also, the final form of the report will correct any inaccuracies pointed out by the partners and will eliminate any unnecessary overlapping of subject matter".

Also in that letter, we added that "since completing the working draft of this evaluation, we have been involved in carrying out the British Columbia component of the national evaluation of the NIP for CMHC, Head Office. The results of the field research involved in the evaluation of NIP have provided us with some different insights and perspectives on the Strathcona experience. We would be happy to discuss these at the meeting, as they will have some
effect on our thinking for the final report of the Strathcona Evaluation". Since the Strathcona Evaluation was supposed to have been completed by the summer of 1976 at least, we had begun negotiations with CMHC about doing the evaluation of the NIP program in British Columbia during the summer and into the fall. We had already designed it, put it into implementation, had all kinds of meetings, and had the printed report ready by January 1977. Obviously some evaluation studies could be completed within anticipated time limits.

Learning quickly, Laurie Hurwitz and I prepared ourselves thoroughly for the review meeting. We went through all chapters of the draft and were prepared to shorten, edit, and delete a certain number of pages from each. We decided that our "fallback position" would be the elimination of all the descriptive chapters, but we refused to change or soften any part of the economic analysis of the project, its budget or impact. We also agreed to give in on one critical issue, the question of naming of names. Geach as well as the provincial representative were adamant that we drop the names of staff members and committee members from the text of the report. Their rationale was unclear; however, it was agreed that names would be omitted. In retrospect, of course, one has to question whether elected or appointed government officials have the right to that kind of privacy when the spending of public money is involved. However there are no "sunshine laws" in Canada as in the United States, and civil servants can remove names from a policy document under the guise of "objectivity", or simply as a means of protecting themselves.

When the meeting came to order on February 23rd in the SPOTA "boardroom" the room was filled and there were double rows of chairs at the table. The government representatives sat at one end of the table, and Laurie Hurwitz and I, with all the other members of the study team, were in chairs at the opposite end. We had not known that the entire team (work base) would turn up
and were enormously pleased by the show of strength and support.

The meeting began by addressing the problem of the length and repetitiveness of the draft. Many of the SRC members, both government and SPOTA representatives, suggested that we drop the initial chapters that gave background on the Strathcona area, the history of urban renewal efforts in the community, and the description of the day-to-day operation of the rehabilitation program. Several considered this type of material unnecessary in an evaluation of the project. There was a suggestion that anybody who was interested in the project would know all about these things and it was not useful to keep them in. However, we quickly received interesting and unexpected support from a new actor on the scene, Doug Ausman, a recent recruit to the British Columbia Provincial Housing Department from the Housing Department in Alberta. He felt that it was critical to leave in these sorts of details. He stressed that many of the reports that he had read left out such information. Therefore they were useless for people coming later who were interested in understanding some of the characteristics of the community in which the program was going on and the general ethnographic details of how a program was being implemented on a day-to-day basis. Being a new presence on the committee and coming with the caché of having been the "golden boy" in the Housing Department in Alberta, his comments seemed to overcome the objections of the government members of the committee as well as the SPOTA representatives, so we wound up agreeing to retain those background chapters.

The next issue which was placed on the agenda by us was the question of how to define the activities labelled "urban renewal" since Geach had insisted that we use "redevelopment" to describe the program that had preceded the SRP. Essentially there seemed to be some confusion on all our parts as to what the use of the term urban renewal meant, since both my partner's experience in Toronto and my experience in the United States had been that urban renewal
referred to clearance. It was hard for us to use the euphemism "redevelopment" for what was essentially a program of expropriation and clearance of housing inhabited by low income house owners and renters. Therefore the quibbles about the use of the term urban renewal in the report reflected to us some critical differences in ideological perspective on the whole question of housing policy.

There appeared to be a lot of distress about the statistics regarding the enormous amount of money that had gone into underground services in comparison with the amount of money that had gone into rehabilitation of homes. Early in the evaluation research members of the study team found it strange that there had been little monitoring of the program as it was going on. Only at this meeting did most of the other actors realize that those aspects of the Strathcona Rehabilitation Program which included a significant input from community groups (the housing rehabilitation component and the recreation improvements) involved less than 20 per cent of the total SRP budget, whereas those items where control remained almost exclusively or largely in the hands of government equalled 65 per cent of the total budget. An example of those are the sums spent on the original hard services. This conclusion, which the report documented rather succinctly and effectively, made many committee members extremely uncomfortable. While the community representatives were clearly interested in these results, the government representatives were not particularly pleased with these figures or our interpretation. They did not want to know this kind of information. Apparently they wanted to be kept in the haze of good feeling that their consensus model of decision making tended to reinforce. (I must say in defence of the representatives at the provincial level, that they had a much better insight into what was going on while the project was happening and had argued that certain kinds of restrictions should be loosened).
It would seem to me, in retrospect, that much of the distress with our report and the eagerness to "disassemble" it was caused by statements that were critical of the planning and implementation of the rehabilitation program, the lack of monitoring, and the narrowness of focus in much of the decision making that was done. Many government people wanted to claim Strathcona as their own private success, in that it appeared to disturb them that the budget spending patterns did not reflect what they thought was happening with regard to the program. It is unclear at this point whether they did not want to know, or they knew and did not want to make changes, or just what happened. Nevertheless, substantial portions of that money in the Strathcona project were documented as having gone to underground services and not to housing rehabilitation or community recreation facilities, which were supposed to be the things being done for the people of the neighbourhood.

The meeting ended with a careful show of agreement among all parties, and a detailed list of revisions and deletions which Ptarmigan was to implement in the construction of the final draft of the evaluation report. But municipal displeasure continued, as the following excerpt from a letter from Geach to Hurwitz on May 19th 1977 illustrates:

"With regard to our telephone conversation of May 17th 1977, I am concerned at the latest estimate of approximately the end of June as the date for completion and delivery of the Study. This will be more than four months after the Strathcona Rehabilitation Committee asked you to edit and complete the Study. It is appreciated that the editing the Committee called for involved a significant amount of re-organizing and re-writing.

One of the effects of this latest estimated completion date will be that the review and reporting on the final document at the City level is likely to be further delayed. I will be unable to give it my personal attention until some time after that date, depending upon other priorities. I think it unlikely that other staff will be available at that date to deal with it. Another potentially more adverse effect may be in the area of reliability of forecasts versus actual delivery date."
While I am well aware of the many reasons for the extended time required to complete the Study, and am aware also of similar delays on Stage I before I assumed the chairmanship of the Strathcona Rehabilitation Committee, I believe that one of the results will be that the relevance of the Study will be significantly diminished. This is particularly unfortunate in view of the Strathcona Rehabilitation Committee's generally favourable reaction to the work to date and to the draft which it considered in February of this year.

The message of this letter appeared to be a warning that the delays and late delivery of the final report would have a series of negative consequences. Not the least of these, he believes, is that "the relevance of the Study will be significantly diminished." However, the members of the study team immediately recognized that the remark reflected a certain paradox of the policy process; that is, that this pilot project had continued to be evaluated long after the related large-scale government-funded program, NIP/RRAP, had itself been implemented in the field for approximately four years and had undergone evaluation in a process that was designed to meet a 1978 deadline for possible legislative extension.

The evaluation process neared its completion. Other than a letter on June 9th, which discusses responsibility for paying for ninety extra copies of the final report and adding a Chinese version of the report summary, the only remaining pieces of correspondence regard the final report itself. We submitted 100 copies of the final report on July 20th 1977 along with 300 copies of the 10 page summary. And on August 24th there is a letter to Mr. (sic) L. Hurwitz telling him that Vancouver City Council in its meeting on August 23rd, received a report from the Director of Planning which indicated that "the study is a thorough and critical review of the project. It makes numerous recommendations which form a valuable guide and checklist for those involved in similar projects."
4.3 The SRP Stage II Evaluation: Analysis of the Relations Between Structural Context, Negotiation Context, and Policy Consequences

To begin this analysis I want to summarize particular components of the structural context of the SRP, the more or less stable elements of that organizational order which provided both obstacles and resources in the negotiation context of the program evaluation study. At the most abstract level of that structural context we can recognize beginnings of a gradual shift in housing policy being articulated by government representatives and community groups across Canada during the late 1960's and early 1970's. This was a complete shift in overall housing policy: from a production orientation to an orientation stressing conservation of existing housing stock; from relatively large-scale demolition and dense new construction to selective removal of uninhabitable dwellings along with rehabilitation and low density infill housing; and from a general exclusion of neighbourhood amenities to an emphasis on the upgrading of both above and below ground facilities and services.

Such a policy shift represented an increasing federal presence in the funding and delivery of programs concerned with social housing and the needs of particular target groups outside the traditional private housing market. In the case of the SRP government cost sharing was split 50 per cent federal, 25 per cent provincial and 25 per cent municipal. The target group of eligible participants was defined in terms of a sliding financial scale which gave full grants of $3000 to those with incomes of $6000 and under and reduced the grant level to $0 at incomes of $10,000 and over (Levitan and Hurwitz 1977:152). In addition, the new policy was supposed to increase levels of citizen participation in the planning, implementation and evaluation of newly developed housing programs. Thus, the SRP, as a pilot project in the testing of this new policy included citizen input in various aspects of project management. However, as noted earlier, during the crucial decision-
making which preceded the formal granting of the SRP contract, SPOTA did not have the necessary professional/technical representation to argue strongly for program provisions that would give community-determined needs greater priority and budget protection. Furthermore, the consensus model of decision making, initiated by the first Chairman of the SRC, entailed an implicit "human relations" approach to management (Kanter 1975:46) which emphasized informal social relations and management rationality. Thus, the essential imbalance of power, skills and resources was never directly confronted by the four parties on the SRC.

However, other aspects of the SRC's position in representing the community's interests were seen by Project participants as more positive examples of what could be achieved by the partnership. Thus, despite the existing legislative mandate that limited SRC spending to specific types of physical rehabilitation in the Strathcona area, review of the Committee's work in managing the Project suggested that members expanded that mandate to assume a role as the unofficial planning committee for Strathcona (Levitan and Hurwitz 1977:28). It coordinated both the housing rehabilitation and neighbourhood improvement components of the Project, and made recommendations to the City on several development issues facing the area. Given the initial community hostility to government and government-sponsored programs (Levitan and Hurwitz 1977:7), the emergence of the SRC as local area planner/advocate contributed to a significant shift in the experience and attitudes of local Strathcona residents. It was in this atmosphere of conditional success of a new government program (as defined by government representatives) and a selective "thawing" of previous community hostility toward all government policy initiatives that the negotiations for accomplishing the Stage II Evaluation of the SRP were undertaken.
Basic Areas of Negotiation

The initial contractual arrangement which specified the provisions of the SRP had also provided for formal representation of the local community (through the SPOTA executive) and the three levels of government on the management committee which was to oversee project operation. This decision to include SPOTA on the SRC was treated as an innovative move in the policy process, an example of government responsiveness to the politically expressed need for citizen participation in public policy and program delivery (Draper 1971). This structural arrangement had a direct effect on several aspects of the negotiation context of the program evaluation research. First of all, it produced a very large "cast of characters" who all became engaged in the various phases of the evaluation process leading up to the critical review meeting of 23rd February 1977. During those protracted negotiations which may be said to have peaked on that date and gradually subsided to the delivery of the final report in July 1977, the nine members of the study team had formal and informal dealings with over eighty individuals (not including the fifty survey participants), from SPOTA, the three levels of government, other community residents, staff of social service agencies operating in the area, the Board and staff of the Strathcona Community Centre, several City departments, staff at the Ministry of Municipal Affairs in Victoria, a range of architects and planners who had been involved at various stages in the SRP, and the entire Site Office staff who worked on the project from 1971 to the end of 1974.

In delineating the major actors in the evaluation study, I have to emphasize the representatives of the four partners to the original program agreement. Although the federal government, through three successive representatives from CMHC, had contributed the major share of the funding of the project, they exercised little overt interest or control over the
implementation of the project (except for the grant/loan formula) and minimal concern over the two stage evaluation process. By contrast, the provincial government, through their three successive representatives from the Department of Housing, had exercised considerable interest in the project and appeared to be most knowledgeable about the impact of its structure as well as day-to-day operations. This level of knowledge and interest carried over to the assistance we received in doing the research (especially in interviews and file searches) and the support we ultimately received regarding the content to be included in the final draft of the evaluation study.

Clearly, we had the greatest difficulty with the last of the three representatives of the City, Mr. Geach. Interviews with other civil servants suggested that the preceding City representatives had what might be termed a more accommodating "theory of negotiation". The difficulties we, as a study team, had with the City Planning Department representative can be interpreted as deriving from the tension between researchers with an investigative outlook and the civil servant who is heavily invested in a procedural outlook (Cronbach 1980:331). Similarly, we might locate this civil servant's approach as reflecting the logic of official interpretation discussed and developed by Don Handelman (Handelman and Leyton 1978) and ultimately based in a set of bureaucratic perceptions and interpretations which are designed to transcend the details of project implementation "in the field".

With respect to the fourth partner, SPOTA, we found their representatives the most fully informed, through their day-to-day involvement in the community and the program, and their ongoing communication with all government representatives including MLA's and MP's. SPOTA members were generous in their contributions of time and energy in the research process and appeared to perceive us as allies and sources of advice and technical assistance. Our renting of vacant space in the SPOTA building clearly assisted them at a time when they
were between grants and money was in short supply. This is an example of the type of exchange relationship (Hessler and New 1972; Hessler, New and May 1979) which developed between the study team and members of the SPOTA executive.

In examining the question of the organizational placement of the program evaluation work it is clear that the Stage II study was designed to be external to program operation both organizationally and in terms of timing. The site office had closed and all housing rehabilitation work had been completed in 1974. However, throughout the period of the Stage II Evaluation the SRC continued to meet on a monthly basis to arrange for completion of neighbourhood improvements and spending of previously uncommitted funds. Laurie Hurwitz and I attended those meetings and were able to obtain substantial field data on the mode of operation of the SRC, the limits to what could be negotiated with Geach as Chairman, and the considerable time and energy expended by committee members in reaching rational, mutually agreeable resolutions of relatively trivial (in budget terms) problems connected with landscaping and the equipment for play areas.

The Stage I Evaluation which was also external to the project had begun while the Site Office was open and all project activities were in full operation. However, their final evaluation report (Bell and Moore 1975) was something of a disappointment to SRC members and the community alike (Levitan, field notes January 1976). It was essentially a brief managerial review which offered partial data (expected because the project was not yet completed), little or no interpretation of that data other than a listing of types of items rehabilitated on houses/properties, and no analysis of potential program impact. Members of our study team rejected what we considered the "minimalist approach" of the Stage I Evaluation (they were paid in full however).

We saw our approach as requiring a much more extensive set of methods and a significantly expanded data base. Initially our research strategy was
essentially a consensual approach to the task of evaluation. However, after the initial draft of the financial analysis came in from the economists, and portions of the participant survey were completed that research strategy started to change. We began to employ an approach that was implicitly contentious in that we saw part of our research findings serving as a "moral witness" for the failure of the SRC to honour its commitments to providing adequate funds for the rehabilitation of homes in the Strathcona area (13% of the budget) while sinking a substantial amount ($1.5 million or 38% of the budget) into below ground services (sewer and water lines) which were basically a municipal responsibility and benefitted those who would profit from increased development in the area. In the final draft of the report we summarized the overall pattern of policy formulation, implementation, and resource control on the project in a single table (Levitan and Hurwitz 1977:113) which fully documented our critical analysis of resource expenditures.

Despite the increasingly critical nature of our research stance, none of the four partners attempted to exercise control or interfere in the development of the conceptual framework or methodology once the proposal was accepted and Ptarmigan Planning was officially hired for the evaluation. Moreover, within the study team (work base) we operated as an extremely egalitarian work group. There was excellent communication and a high level of trust and "camaraderie" among the nine main members of the team. Often there are serious differences in conceptualization and work-style among members of interdisciplinary work teams (Vivelo 1980). However, the diversity of skills and disciplines did not impede communication and, in fact, contributed to the comprehensive approach which was achieved in the final draft of the report.

All the methods of inquiry and data sources used in constructing the evaluation (Levitan and Hurwitz 1977:149) were designed to produce information
on the major components of the project: physical, social, economic and political. There was a particularly good fit between the data that came back on the social planning component, from the physical planning component and from the economic analysis. Thus, the first draft that we assembled did not have to mask any serious contradictions in outlook or content. No doubt it was repetitive and too long -- but we all suffer from that problem, pure and applied researchers alike.

It is difficult to attempt to understand all the respective stakes in these negotiations. The stakes of the evaluation team were mixed. They included meeting the contractual agreement to get the evaluation accomplished and payment involved. But what seemed to be more significant was the chance for members to work together collectively and enjoyably while learning much more about doing evaluations of government programs in a work setting that was both complex and yet accessible. As for the government representatives, in a sense they all wanted to take credit individually for the success of the Strathcona Project. And SPOTA wanted to take credit for the project since, from their perspective, had it not been for their "last ditch" organizational efforts there would have been no rehabilitation project and Vancouver would have had its counterpart to Halifax's Africville.

Most of the transactions during the initial phases of the evaluation were relatively visible, and the members of the evaluation team worked to maintain that visibility. We tended to be somewhat naive about the hidden agendas of a few of the negotiators. In particular, we underestimated the power of the Chairman of the SRC, the Assistant Director of Planning, who appeared to have an unswerving professional and personal investment in slum removal as his preferred intervention strategy (Warren 1971). Thus, the pressures which emerged in the months between receiving and reviewing the preliminary draft of the evaluation document forced several members of the
study team to reassess that strategy of overtness and visibility. We shifted to a more manipulative stance, which centered on a plan and a "script" for salvaging all that we valued most from the preliminary draft while negotiating at the review meeting. We were united on what compromises we were willing to make and what sections we were willing to jettison in working toward a mutually acceptable format for the final draft.

In retrospect it is evident that neither the conceptual framework nor our methodological strategies became negotiable issues as part of the evaluation process. We did not have conceptual or methodological confrontations among members of the research team where a spirit of cooperation predominated. Essentially the members of the SRC never directly confronted the legitimacy and competence of the evaluation team. Rather, it was the vocabulary and content of our interpretations of the research results which became the focus of controversy in the draft review meeting. And the sudden "switching of sides" of the newly appointed and highly regarded representative from the Department of Housing tended to blunt the force of the criticisms.

The assembling and production of the final draft of the Stage II Evaluation was in many ways an anti-climax in this long series of discussions, lobbying sessions, interviews and meetings which had continued for nearly two years, punctuated by superb eight-course Chinese dinners. At no point did we really consider discontinuing our work, but after the review meeting the final report began to assume the shape of an "albatross". We had, by then, completed the British Columbia component of the National Evaluation of NIP/RRAP (the next case study) and began to understand that there was little chance that our recommendations from the SRP Evaluation would find a hearing in Ottawa.

The final report was an attractive, 160 page, legal size document which included many maps, diagrams and photographs to break up the text. The
reactions included: "Good Stock" (meaning the paper), the SRC Chairman; "It was easy to find the figures on what was done and how much it cost", the British Columbia Regional Director of CMHC; "It shows how unique Strathcona is as a community", a previous Chairman of the SRC. Other civil servants asked "How could it be generalized to other communities, considering how unusual Strathcona is?" and "Why would there be any interest from other communities?" There were several comments that involved what I consider 'discounting strategies', i.e., responses that pointed to the idiosyncracies of the Strathcona community, that claimed to be positive and yet served as a way of discounting the generalizability of results from the Strathcona Project. It became apparent that certain members of the SRC and other civil servants were still invested in the myth that somehow the Chinese community had "ripped off" the government (Levitan, field notes February 1977). Our investment analysis which documented that homeowners had carried 9.6% of the total SRP costs (Levitan and Hurwitz 1977:58), while the City of Vancouver carried only 19.6% (the province 23.6% and the federal government 47.2%) certainly contradicted that myth. The perversity of that political myth was all the more significant since over two-thirds of those homeowners (by no means all of them Chinese) had family incomes under $6000 per year, and part of their savings had assisted in subsidizing municipal services which could ultimately destroy the livability of their own community through encouraging private redevelopment in the Strathcona area.

Essentially, the SRP Stage II Evaluation emerged from small-scale controversy into a planning/housing "limbo". Its content and recommendations were never to play any role in shaping future programs of neighbourhood rehabilitation and/or housing rehabilitation. Only the Acting Director of the School of Community and Regional Planning at UBC found it "an important contribution to the literature", hoped it would be considered seriously, and
indicated that it would be used in the classroom. Up until 1980 I continued to receive calls for information on Strathcona from students in the Planning and Architecture Schools. But those have ceased, and housing rehabilitation (as "home renovation") has become a profit-making part of the private sector of the economy.

Summary and Implications for Knowledge Production and Use

In summarizing the relations among these various negotiators who were the dominant actors in the Stage II Evaluation, it is somewhat difficult to characterize in a single phrase the social order which evolved. There were many varying levels of cooperation and antagonism which had been part of the structural context at the beginning of the SRP. These had continued into the operation of the project and underwent little transformation in the process of our (Ptarmigan Planning's) involvement as program evaluators.

Initially, one representative of the municipal government (Director of Social Planning) persuaded the other parties to the SRP agreement to use a form of decision making (the consensus model) which was perceived as a means to minimize the conflict and antagonism brought to the situation. Clearly, the consensus model only masked the conflict between SPOTA and the City. Ultimately, some of that same conflict surfaced in relations between the last City representative (indirectly supported by his Engineering Department) and the research team in the Stage II Evaluation.

Otherwise, there was a general level of cooperation among the two other government representatives, the SPOTA executive and rank-and-file, and members of the study team. Furthermore, as frequently indicated, the study team itself was extremely cooperative and always presented a "united front" in dealings with civil servants who were unjustly critical of the SRP. The
relatively egalitarian organization of the interdisciplinary research group evolved into a "durable web of commitments" (Strauss 1978:106) which provided the support system for development of a comprehensive, sophisticated evaluation study.

Although the SRP was originally characterized as a new, ambiguous, and risky experience among the four partners, ultimately the project settled down to a routinized pattern of combined cooperation and continued conflict for any and all negotiators. Furthermore, as timing alone would suggest, the evaluations (Stage I and II) were never designed to have any subsequent impact on the plans and activities of government decision makers concerned with the next phase in Canadian Housing Policy. Our research work was carried on far away from the center of change in the policy intervention process, and we had no network of personal connections which might have enhanced its "reception" by the policy makers.

However, looking more closely at the negotiations which developed in the course of the evaluation process there are certain generalizations that may be drawn from the research experience. These included characteristics of the negotiations which became evident early in the evaluation and continued to affect the later phases of the evaluation process. First, there were a very large number of actors engaged in the evaluation process, policy makers, civil servants, community representatives, project participants, architects and planners, and a nine person study team. Its seems evident that the four-party agreement to manage the SRP and the generally high level of interest among most representatives to the SRC encouraged a pattern of widespread involvement in anything affecting the project, including the project evaluation. Second, there was the constant shifting from formal to informal styles of communicating and conducting business. This tended to be confusing at times since the evaluation team members were not fully aware of the implicit rules
for making that shift. In our observations of the SRC meetings we gradually learned some of the hidden agendas of the various government representatives, understood that the SPOTA leadership would generally find ways around bureaucratic obstacles and never used a confrontational approach, and finally grasped the conservative, careful and rational nature of all SRC discussions and deliberation. Yet all that changed in the open, gossipy, high spirited interaction of eight course dinners and community celebrations where individuals spoke freely about community conflicts, serious disagreements between SPOTA and the City on specific development issues, and the overall "tight-fistedness" of CMHC.

The large numbers of actors and the shifting formality/informality of the interaction process contributed to the protracted nature of the evaluation negotiations. It took nearly six months for study team members to construct a negotiating strategy and push through to that review meeting in February 1977 where we finally reached some agreement on the style and content of the final report.

There are several aspects of the negotiations in the case study which have implications for the construction of additional models of knowledge production and use for policy research. In research settings characterized by a large number of actors, it may become essential that policy researchers quickly study and identify the political, economic and organizational stakes of the key actors before formulating a general research strategy and the model and methods for fieldwork. In addition, members of the study team may find it important to carefully discuss and assess the implications of the research strategy they plan to take. Whether explicit or implicit, any research strategy has certain identifiable assets and liabilities in specific research settings. The planning for ultimate research use often begins at this stage with an awareness of constituencies to be served and the
identification of policy opponents.

The nature of the relationship between evaluators and program funders and managers has assumed a central place in the literature on program evaluation. However, the dichotomous external/internal approach tended to overgeneralize what is often a complex set of relationships between program evaluators and several other actors in the research setting. Although in this case study the organizational placement of the evaluation work was predetermined (bureaucratically required as an "external evaluation") some aspects of an "exchange relationship" with several key actors contributed to the depth, and accuracy of the evaluation report (Rowbottom 1977). However, this did not guarantee acceptance of the research findings. Similarly, the use of very traditional quantitative methods (e.g., investment analysis, analysis of financial statements) may be powerful tools for debunking organizational myths. However, where policy makers and civil servants prefer those myths, some form of bureaucratic strategy may be employed to "bury" such critical findings (Dennis and Fish 1972).

In general, the analysis of this case study suggests that careful technical and design work done in evaluation research studies are not sufficient to secure a "positive reception" for research results. Any alternative models of knowledge production designed to be used in policy intervention must take into account the structural context of the research setting, including the "value screens" of policy-makers and other research users.
4.4 Case: Evaluation of the Neighbourhood Improvement Program Delivery System in British Columbia

The second case study focuses on a program evaluation carried out for CMHC in 1976-1977 by a combined study team representing the Community Planning Association of Canada (British Columbia Chapter) and Ptarmigan Planning Associates. The evaluation was one of a series carried out by consultants for the Neighbourhood and Residential Improvement Division (NRID) of CMHC. The evaluation report analyzed the Neighbourhood Improvement Program (NIP) Delivery System in British Columbia, emphasizing the organizational aspects of that system and examining, in some detail, the processes developed for gaining citizen participation in program planning and implementation.

As with the Strathcona Project, however, I offer here a description and analysis of the actors, issues and negotiations which were central to the production and consequences of that report. In this first section of the case study I emphasize those areas in the evaluation process which were perceived as the most significant problems to members of the evaluation research team and were documented in my own field notes, written reports or a recent interview with my former partner in Ptarmigan Planning, Laurie Hurwitz. In Section 4.5, as in 4.3, the analysis represents my own perspective on the evaluation research process and is guided by the analytical framework developed in Chapter 3.

The Structural Context of the NIP/RRAP Programs

In 1973, Parliament passed legislation setting up NIP and the Residential Rehabilitation Assistance Program (RRAP). The two programs, separately administered, but closely related were designed to: "conserve and rehabilitate housing stock, add or rehabilitate required social and recreational amenities
or municipal services; remove blighting land use; and promote the maintenance of the neighbourhood after the NIP project terminates (NIP Operations Handbook 1975: A-2). CMHC had recognized:

the fact that it is more desirable in both social and economic terms, to conserve and maintain the desirable qualities of a residential neighbourhood before it deteriorates so badly that major urban surgery is required.

The NIP program was to parallel rather closely certain aspects of the American Model Cities Program, inaugurated in 1966 to rebuild or revitalize slum areas. Just as with Model Cities, the NIP program was to be a demonstration program which "does not seek to provide unlimited resources to do everything that may be considered desirable". Instead it required communities to do their planning and implementation within the terms of "known and predefined resources". This initial financial planning was then to be followed by an ongoing implementation process that was to include resident participation. It was hoped that such a neighbourhood plan would develop:

an approach to the integrated and local delivery of social, health, education and recreational services, such as legal aid, family counselling services, and the like.

Both the NIP and RRAP programs started in cities and towns across Canada in 1975. From then until the end of 1978 the federal level of government paid the largest share in the funding of the programs (approximately 50%) and received the most substantial political "credit" for the effectiveness of the programs (Mitchell, personal communication 1981). Toward the end of the 1970's there was another shift in the balance of power between the federal and provincial levels resulting in further changes in the National Housing Act. Both the NIP and RRAP programs were reduced, and the federal government arranged to deliver part of those budget allocations into provincial hands through "global agreements" which do not specify rules for spending, just the requirement that it be spent on housing. Consequently the provinces
now have the right to provide funding for some types of social housing, while the remaining social housing is delivered through private sources (e.g., ethnic/community groups are given funds to build housing for the elderly). Unfortunately, the resources for neighbourhood improvement, which were also turned over to the provinces in 1978 (but with a much smaller budget than NIP), ended in 1981 (Mitchell, personal communication, 1981).

**NIP and RRAP: Citizen Evaluations**

Various conceptual/procedural issues were at stake in the delivery of the NIP program across Canada; because they affected our evaluation of that program here in British Columbia, they merit some attention as another aspect of structural context. Specifically, the NIP program stipulated that the residents of a NIP area must be involved in the physical and financial planning for their neighbourhood. Since there had been limited experience in citizen participation in any such processes in the past, a pilot program to explore participation was launched in 1973. Under this program, eight projects were funded in various cities across the country to determine the role of citizens in the NIP planning process. (There was no promise that any of these areas would be later designated as NIP areas). While it was recognized that each area would have its own identity, it was also felt that there would be some similarities between the groups which would assist governments in establishing guidelines for citizen participation and would also be helpful to other groups who would be going through the same process under the NIP program (Bean... 1975).

At a conference to evaluate these NIP pilot projects, held in Vancouver in May 1974, the various citizen groups involved presented their evaluations and recommendations. With regard to the Residential Rehabilitation Assistance
Program, they agreed that the program income limits should be more sensitive to regional differences and be raised to somewhere between $7000 and $12000 depending upon the area. They felt it was "unfair to restrict housing rehabilitation to areas that are already receiving a considerable amount of money under the NIP program." They approved of the continued separation of NIP and RRAP not for purposes of administrative simplicity, but because the separation was interpreted as a fairer means of distributing the limited resources available. Second, the citizen groups recommended redistribution of authority within the CMHC bureaucracy, increasing the decision-making authority at the lower levels. "Local officials", they felt, "should be given more discretionary powers. It is frustrating for groups to have to deal with representatives who cannot answer questions directly, but constantly have to return to their offices and consult others and use up more time before getting answers."

Third, the citizen groups felt that neither they nor the municipal government received adequate information and guidance. The resulting delays (often several months) in implementation of NIP had caused nothing but frustration for all concerned. It was recommended that "all governments and citizen groups have to work together to familiarize themselves with the mechanics of the program to avoid delays in the future and thus to strengthen the relationships between all concerned."

The pilot projects had revealed the diversity of settings in which the NIP program would be implemented, the varying roles of the citizen groups within these settings, and the differing vested interests which would be involved in the programs. It had become obvious that there was "often a problem with emphasis between the two groups of people most involved in implementation of the two programs." Planners were most interested in, and therefore placed most emphasis on, the NIP program and general community services. The citizens were often more interested in RRAP because this program would have the most direct effect on them. Conference delegates
therefore recommended that, before implementation of either of the programs, "a clear agreement should be reached between the groups and their respective municipal governments as to what the role of each should be. If a mechanism is set up in the beginning as to how decisions are to be reached, it will save a great deal of trouble when differences arise". They also urged citizen groups to work "towards developing good negotiation skills, presenting their case clearly, lobbying, good spokesmen, so that they can have strong delegations to present their case to government officials."

Finally, the citizen groups expressed their awareness of the problems involved in making a distinction between physical and social rehabilitation. As the representatives from Inglewood-Ramsay stated: "urban rehabilitation as opposed to renewal must be recognized as a social, not just physical, process. People, not buildings, should be our foremost concern. Unfortunately, it seems thus far the physical side has been emphasized'.

Thus, two years before the 'professionals' were invited to evaluate the NIP and RRAP programs, lay people had produced a sophisticated critical analysis of the conceptual and operational issues central to the success or failure of the programs. They were cognizant of the difficulties related to the distribution of resources and authority implicit in the rules and procedures governing the programs. They warned of the problems connected with inadequate information and delays in implementation. They had a clear notion of the enormous variety in the settings and relationships from one program site to another and an understanding of the differing vested interests in the success of the program components.

The British Columbia Evaluation: A Brief Chronology

Several aspects of the overall evaluation context of NIP and RRAP had
consequences for the shape and content of the evaluation strategy developed in the British Columbia context. These included properties of the program to be evaluated, as well as properties of the strategies for evaluation which had already begun to be negotiated by various actors in Ottawa and Toronto before the representatives of Ptarmigan Planning had entered the process. The NIP/RRAP program in British Columbia was part of a national program which had already become somewhat standardized, organized to follow a relatively conventional hierarchical approach to decision-making and designed to include community input only at the local level of neighbourhood improvement. The national strategy for evaluating the program was coordinated by a CMHC representative in Ottawa and directed by a sociologist in Toronto. The British Columbia evaluation was to be accomplished at the local level by CPAC (B.C.) and Ptarmigan.

In contrast to the Stage II Evaluation of the Strathcona project, the pre-contract negotiations for the NIP Evaluation were brief and informal. As a result of our work in Strathcona and our familiarity with the results of the 1974 conference, the principals of Ptarmigan Planning (Laurie Hurwitz and I) were well informed of the problems that NIP was encountering in the field. We decided to initiate contact and, early in 1976, we sent a letter to the Research Director of the Neighbourhood and Residential Improvement Division of CMHC to inquire about plans for evaluating NIP and RRAP and to offer our services. Several phone calls later we were offered part of the national evaluation study. There was no proposal call for the evaluation study and we were never formally interviewed for the contract. In fact, I have never found any letter of agreement between ourselves and CMHC.

Thus, from the beginning, ambiguity clouded the negotiation process. For example, in our early conversations with Nils Larson, research director of NRID, he offered us the British Columbia component of the national evaluation
study. Then, in a further series of telephone calls in May 1976, he told us that Ptarmigan Planning had been awarded the contract jointly with the British Columbia Branch, Community Planning Association of Canada (CPAC). Larson assured us that any contractual problems could be resolved at our first set of meetings in Vancouver in August. But, prior to that meeting, we discovered that NRID had arranged for CPAC to be the major contractor, do the contract and financial administration, and provide clerical support. Despite his assurances, it was never made clear to us why the original contract was divided up, and who had the power to make these decisions.

In August, the five members of the study team (three from Ptarmigan and two from CPAC) met with Larson, representing CMHC; the national research coordinator, a sociologist from the Ontario Institute for Studies in Education (OISE); and the CMHC Regional Office representative of the NIP program in British Columbia. For two days we discussed our roles in the overall evaluation strategy and analyzed the proposed methodology designed by the national coordinator. The following document sets out the context within which our specific set of four case studies was to operate.

August 17th, 1976

Current Strategy for NIP/RRAP Program Evaluation

NIP/RRAP program evaluation is being undertaken for two main purposes. At one level, evaluation will provide the basis upon which adjustments can be made to the existing programs, to increase their effectiveness within the scope of existing program objectives. At a second level, evaluation will provide a sound base for the preparation of new and/or revised legislation, a process which is wider in scope and may involve the recasting of program objectives to attain the broader policy objectives.

The latter point is especially relevant to the NIP program, since the NIP legislation expires in March 1978. Working backwards, this would imply a submission to Cabinet in June 1977. This date will necessitate a policy submission to CMHC Management in April 1977 which in turn necessitates a Division paper by March 1st, 1977. It is in the light of this (approximate) schedule that the Division has prepared a Part V research plan which involves three elements:
1. An evaluation of the NIP Delivery System, through a series of 21 case studies. The research design phase of this project is almost complete, and the case study phase has been approved in principle, for completion on December 1st 1976.

2. An evaluation of the Social and Economic Impact of NIP and RRAP. This proposal will also involve the analysis of a series of case study areas, with additional study of a number of control areas. One problem integral to this study is the fact that much of the NIP/RRAP impact (on mobility, land and housing values, etc.) will be apparent only in the long term, so a study conducted in 1976 would only offer an interim assessment -- this is nevertheless seen as useful enough to proceed with the study, to be supplemented at a later date.

3. A NIP/RRAP Need and Demand Study, to estimate the need and demand for both programs on a provincial level of aggregation, so that CMHC can assess total need. This proposal will be submitted to Management in August.

In addition to this Part V activity, there are a number of complimentary in-house tasks which will be undertaken.

1. The existing monitoring system for NIP and RRAP will provide a wide and shallow data base for policy development, as compared to the less extensive but deeper analysis resulting from case studies. (Note that there are 21 case studies, but over 400 NIP projects will be generated by 1978).

2. An interdivisional policy working group will be established, to discuss inter-program policies related to the NIP and RRAP target area types.

3. Provincial positions on revisions to NIP legislation will be sought out during the period September - December 1976. The provincial input will form part of the evaluation study material.

All these evaluation activities will result in an intensive internal (NRID) debate and report formulation during the period December 1976 to March 1977.

The work proceeded with dispatch. (The Strathcona project was still in the process of being evaluated, and there was a great deal of overlap). We started work on the NIP evaluation interview schedule on August 20th, and continued to work on it into September. We worked on the RRAP interview schedule in the middle of September but when it was found to be a much too expensive process to do the RRAP evaluation work on the schedule ceased. Field work on the evaluation of the NIP program continued from September until November. We completed the NIP evaluation report in December 1976; it was
sent to Ottawa and typed there in January 1977.

The Evaluation Process: Methodological Conflict

Although the study proceeded with dispatch and apparent efficiency, discord and antagonism characterized the entire process, typified perhaps by the fact that several letters sent to us from Larson were addressed to the 'West Coast Mafia'. A series of conceptual and methodological difficulties emerged in the frustrating encounters between ourselves and Larson who represented CMHC, the organization funding the evaluation study, and the national coordinator, who attempted to control the study design.

Three major problems developed with regard to the methodology designed by the national coordinator. These concerned: (1) the lack of clear definition of the research problem; (2) differences with regard to the purposes and structuring of the interview schedule; and (3) the limited benefits and high costs of the RRAP/Impact study design. These are discussed below, along with ways in which the team worked to overcome the problems.

Our first major difficulty was the lack of a clear definition of the research problem and the necessary set of operational definitions which would serve as the framework for the field research. Both the draft and the final study designs identified the problems to be researched in vague, general terms, or alternatively, in terms of excessive, complex detail, focussed on the problems of data collection.

We should emphasize at this time that we are not referring to a lack of definition of the purpose for the evaluation. We feel that CMHC has been perfectly clear in what it hoped to obtain from the evaluation. CMHC was interested in two areas:
1. With the legislation for NIP "running out" in 1978, CMHC was interested in knowing if that legislation should be continued and, if so, were there any desirable changes as noted by the various participants who were involved in NIP;

2. The effectiveness of the delivery system, and ways in which it would be made more effective if the program were to continue.

Thus, it was stated that the research purpose was to evaluate the overall operation of the program, and not the functioning of any project in particular. But a clear sense of purpose is not a substitute for the more specific research questions which are needed as a focus for the emerging study design. We should also note that numerous, complex charts indicating concerns and areas of data collection do not, in themselves, adequately replace the coherent problem formulation which we consider to be critical in a national evaluation.

Instead, in British Columbia, the study team had to develop a strategy for covering the major theoretical issues raised in the national study design. This took the form of a series of six (and later seven) questions which were derived from the statement of purpose and the needs of CMHC. These served as a specifically articulated framework for the interview schedule, and the major headings for the analysis of the results. These questions were:

1. How were municipalities and NIP areas designated? To what extent did they meet existing eligibility requirements? Were these eligibility requirements reasonable and useful?

2. What were the staffing allocations for NIP? How did these allocations affect the pattern and efficiency of program delivery?

3. What was the nature of the administration of the program? Did this produce any significant positive/negative effects on the delivery of the program?

4. A Maintenance and Occupancy Bylaw is a CMHC requirement. Was such a requirement useful? What effect, if any, has it had on local housing conditions?

5. What has been the nature of the decision-making process? What effect has citizen participation had on the various participants in the decision-making process?
6. What effects have there been, or will there be, on the neighbourhood (and the municipality) because of the NIP program?

These six questions served as the "guideposts" for the NIP evaluation in British Columbia. Shortly after the initial implementation of the methodology, we added a seventh question. This question concerned the financial limitations of RRAP:

7. What effects have the RRAP financial formulas has on take-up? How did RRAP affect NIP delivery?

These seven questions represented the operational definition of the research problem as it was interpreted by the members of the study team.

Our second major problem concerned the interview schedules which were to be utilized in the field research. We had a series of difficulties which had to be overcome before we were able to arrive at a satisfactory schedule of questions which could be used in the field with some ease and effectiveness. These difficulties included: (1) the number of different schedules which had been produced; (2) the strategy used for structuring the various sets of questions -- general/specific, concrete/abstract, etc.; (3) the style and wording of the questions themselves; and (4) the strategy of using replicated questions in order to "test" the accuracy/information level of respondents.

The initial package of field material contained sixteen different interview schedules, to be used on the various designated respondents required in each study area. This number of different schedules would have made any comparative data analysis virtually impossible. We felt that even the most skilled research interviewers would have great difficulty in becoming familiar and comfortable with sixteen different schedules. In addition, we were concerned that the exclusion or inclusion of questions for different respondents was basically prejudicial in that it pre-judged an individual's role or interest in the NIP program.
In order to overcome these problems, the British Columbia team developed a core interview schedule which covered all the basic information needs of the national coordinator. It could be used, with minor modifications, with all respondents in the sample; and it followed the operational framework which we had developed.

Our concern with the structuring of the various sets of questions in the schedule stemmed from a basic methodological difference. The national coordinator wanted a "research design that attempted to move from the general to the specific." His interview schedules also followed that pattern, moving from the general to the specific, the abstract to the concrete. The British Columbia team agreed, based on our collective research experience, that interview schedules elicited better data if they moved from the specific to the general -- where probes were used as an "opening up" for the respondents, rather than a "closing down." The first question from each of the final interview schedules clarifies this difference in approach.

**National Research Schedule:**

1. Could you tell me what you see going on in the Neighbourhood Improvement Program on a day-to-day, week-to-week basis? What kinds of things happen in the NIP Program? How does it affect you and how do you affect it? (Probe for experience, problems, details) (Probe by Stage 4:3:8)

**British Columbia Schedule:**

1. What do you feel is important about this NIP Project?

1a. What planning items were chosen as the most important? (Why?) (What other choices were turned down? -- Why?)

As is evident, the question from the National Research Schedule asks the respondent to mentally go through the whole range of his/her experiences and make a series of generalizations about the program for the interviewer. Only then is the interviewer instructed to probe for a selection of the specific items from which the respondent derived the general remarks. Rather than provide the respondents with an opportunity to elaborate on their
experiences, this method leaves the respondents feeling that they have not provided the expected information.

The question from the British Columbia interview schedule asked the respondent to focus on the project in which he/she was involved and to select that aspect which he/she interpreted as most important. Then the respondent is asked to look at the next level of abstraction, all the planning choices made regarding that specific project, and indicate what the decision-making group considered the most important items they had to attend to in their work. Many respondents then went on to make some general remarks about the Program, but only at the level at which they themselves felt knowledgeable and comfortable. (All provincial and CMHC regional officials were asked to describe the Program in general terms -- but that is specifically their job and the level of their expertise).

This problem regarding the direction of the questioning was again paralleled in the style and wording of the questions themselves. The interview schedules designed by the national coordinator tended to use technical and abstract terms, rather long, complicated sentences. The schedule designed by the British Columbia team stressed the use of simple language, commonplace words, wherever possible, and an uncomplicated format. This was achieved by doing a series of pre-tests (and subsequent modifications) with respondents who were willing to struggle through the ambiguities and not feel threatened by that process. These differences are illustrated by the following questions from each of the final interview schedules.

**National Research Schedule:**

Have the criteria for eligibility and allocation and the guidelines been reasonable, practical and useful?  (Probe)

**British Columbia Research Schedule:**

We would like to ask you some questions about the criteria for a neighbourhood to be eligible for NIP. CMHC have six eligibility
criteria -- I will read each of these criteria and would like your opinion on:

1. whether this criterion is a reasonable (fair) limitation for selecting neighbourhoods, and

2. whether this criterion was useful (helpful) in helping to select this particular neighbourhood.

The interviewer was directed to then hand the following card to respondent:

Figure 1: Card with List of Six Eligibility Criteria

<table>
<thead>
<tr>
<th>1. Residential Land Use</th>
<th>4. Moderate and Low Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasonable</td>
<td>Reasonable</td>
</tr>
<tr>
<td>Useful</td>
<td>Useful</td>
</tr>
<tr>
<td>2. Needs Housing Rehabilitation</td>
<td>5. Deficient Amenities</td>
</tr>
<tr>
<td>Reasonable</td>
<td>Reasonable</td>
</tr>
<tr>
<td>Useful</td>
<td>Useful</td>
</tr>
<tr>
<td>Reasonable</td>
<td>Reasonable</td>
</tr>
<tr>
<td>Useful</td>
<td>Useful</td>
</tr>
</tbody>
</table>

The final difficulty with the national interview schedule centered on a strategy of using replicated questions to "test" the respondents. We considered that this "examination mentality" seriously interfered with establishing good rapport with a respondent. And, secondly, as a respondent and interviewer worked their way through the rather extensive interview schedule, new information was bound to emerge and/or issues be looked at from a different vantage point. This meant that responses might naturally differ, depending upon the context. We made every effort to emphasize the mutuality of the interview experience and did not challenge the competence of the respondent.

Such were the problems we had with the interview schedule and the strategies we developed to overcome them.

After a series of discussions and negotiations, the British Columbia study team decided not to participate in the national study of NIP/RRAP impact. There were a number of reasons for this decision which we outline
briefly as follows:

1. It was never clear how the research would be applied.

2. The impact study would have no "before" data, either subjective or objective, against which impact could be assessed.

3. It was questionable whether comparable objective data would be available for all relevant case study areas. This is especially critical in light of the comment in a letter from CMHC dated October 5th, 1976 in which it was noted that "the variation in local standards and size of municipality may make it difficult or impossible to gather some of the objective data".

4. The "subjective" data requested was time-consuming, and frequently redundant. The new interview schedule would have required an additional call-back after the two to three hour schedule had already been administered. Many of the questions required the respondent to go over information already discussed in the first interview. We felt this to be unreasonable.

5. It was obvious that the "subjective" impact instrument and the "objective" impact instrument were not related to each other. It is questionable whether any of these instruments could determine the difference between the "perceived" reality and something called "objective" reality. This is part of a long and often heated philosophical discussion of whether there is such an identifiable distinction.

6. There was a problem with the language of the interview schedule. It was totally inappropriate for use with RRAP users as originally designed, and would have required another week of work to translate it into usable form and wording.

7. There was no rationale indicated for the windshield survey. It was an unreliable procedure at best, and there seemed to be no purpose in doing it as a measure of "impact".

8. The first Objective Data Format could be extremely time-consuming for any of the project teams to research, but especially so for teams doing larger cities and towns where there would be a large number of non-residential facilities. The data would be very approximate and, in some cases, owners would be unwilling to give information. Public records could hardly be expected to be completely up-to-date.

9. The "clincher", of course, was the other five Objective Data Formats. These were estimated by an economist/land use consultant as requiring about 25 person-days at $300 per day or $10-20,000 per case study for baseline data. The data could be obtained for a municipality or regional district, but in order to get it for a NIP area, it would have to be property by property. This would involve a process of obtaining a list of the necessary legal descriptions, then visits to the land registry (fee for each parcel), assessment authority, municipal records, etc.
Although we did not participate in the national NIP/RRAP Impact study, our interview schedule included questions which measured impact as an integral part of the process of neighbourhood improvement and upgrading of the housing stock.

The conflicts described here reflected two contradictory approaches to evaluation research. The national coordinator appeared to be using a form of experimental, positivist model which insists that the researcher exercise a controlling authority over the whole endeavour. This is Clarence Sherwood's "snarling watchdog" approach to program research (Marris and Rein 1974:253). In contrast, the members of the study team worked from a model of program research that assumes constant change and interpretation within any program context. Furthermore, we shared with Michael Brooks (1965) a sense of the ethical necessity for continuous feedback of research findings. We often found ourselves reaffirming the notion that those persons who delivered and participated in the NIP/RRAP programs were the basic "authorities" on the operation of the programs and had a great deal to offer to any set of outside evaluators.

The controversy between us was not merely concerned with the reasonableness of compromises that might or might not be developed regarding research procedures. "It turned on the relationship between research and action" (Marris and Rein 1974:255). The national coordinator implicitly denied any contradictions between the objectives of scientific research and those of social reform. Any apparent conflict between the demands of theoretically derived research procedures and the political or operational pressures in the field suggested that what went on in the field was slipshod and was not designed to get the "true" picture of how the program was being delivered.

None of the members of the British Columbia study team shared his dedication to getting at "the truth". We had all spent several active/frustrating years implementing programs in various areas (housing, education,
recreation, social work, mental health), and in doing the types of research necessary to plan and evaluate such programs. It had been evident to all of us, in different ways, that the orthodox model of social research was inappropriate for producing the types of information/response that is relevant to programs of social reform where objectives and procedures are constantly in flux. Further, we were well aware that any assessment of what the programmer has done or is doing is totally dependent on the respondent's "locus and theory" regarding social change.

After discarding the contents of several large yellow notebooks of sophisticated interview formats, the British Columbia study team developed an interview schedule which accommodated itself to the requirements of the CMHC research questions but also offered respondents a chance to develop their own assessment of the two programs (NIP/RRAP). We used our collective skills to solve the technical questions and produced an enormous amount of data from respondents. And, in the last round of negotiations, we stood firm in protecting the identities of those who had provided that rich and varied data. As Marris and Rein (1974:254) remark: "in social research you are usually either disreputable or unhelpful."

An Ethnography of Conflict: The Details of Interaction

The following description (Levitan, field notes, September 1976) of one encounter between members of the study team and the NIP/RRAP evaluation management team (Larson and the national coordinator) conveys the quality of the interaction which occurred in our face-to-face dealings. Since we had been used to eating in Chinatown during the SRP evaluation, we often had what is called Dim Sum in one of the Chinatown restaurants. We would all sit around the table, usually six or eight of us, order various dishes of various
kinds of things, tote up the bill and divide it six, seven or eight ways, however many people were there.

During one set of meetings we went to the Golden Crown with Larson and the national coordinator when we were so furious individually and collectively with the latter that we somehow wanted to get back at him. Since he always acted in his own interests we finally had a collective situation in which to get back at him. Michael Goldberg kept ordering the most outrageous types of things for lunch, steamed chicken feet, the kind of "bau" that are stuffed with an unusual soybean paste mixture, all sorts of things that only Michael really cared for. The dishes piled up and we kept reiterating that we all shared in the costs of these things. The coordinator kept saying "Oh no, I am only going to pay for the things that I ate, oh no I have a stack here of the dishes that I ate, I am only going to pay for that."

In the end we verbally battered him with regard to his individualist approach, his lack of collective cooperative behaviour, and finally got him to pay a one-eighth share of the total bill. But we had all been reduced to this incredibly infantile, childish level of coercing him into paying for about three extra dim sum dishes, all of which must have cost him in those days about 75¢. This was the level we were reduced to by his total avoidance and rejection of any legitimate claim on our part to have a right to negotiate with regard to the concepts, the design, the methodology, or anything having to do with the planning and results of the evaluation.

The Fieldwork Process: An Exercise in Research Team Conflict

One might assume that the technical solidarity shown in the conflict with the national coordinator (and indirectly with the NRID research director) was paralleled by a social and emotional solidarity among the study group.
But, in contrast to the web of commitment that built up so strongly in the SRP study team, the NIP/RRAP evaluation involved much more internal conflict.

Of the five members of the study team two of us were sociologists, one was a geographer, one had degrees in both economics and social work, and my partner, Laurie Hurwitz, had degrees in anthropology and planning. Presumably this was a relatively homogenous group, with a major overlap in actual personnel from the Strathcona study. We already knew that there were serious differences in the "value screens" among the five of us; that had also been true in the Strathcona evaluation study. However, in this setting we appeared to be unwilling or unable to grant the legitimacy of one another's points or view and it became difficult to divide up the field work in an arrangement that was mutually acceptable. Possibly the ongoing conflict with the "management" of the evaluation study (the NRID director and the national coordinator) contributed to the kind of work divisiveness that is characteristic in "non-union" shops.

Gradually other ideological and task conflicts began to emerge in our five person study team. These were confrontations which seemed to divide us into two sub-groups and one person in the middle. Unfortunately at the time it appeared to be two "macho" males vs. two "feminist" females, and one "non-traditional" male attempting to mediate. At this distance in time it becomes difficult to support or refute that image. However, it may be argued that some level of professional commitment to the research process among the majority of the team members overcame these interpersonal schisms so that we could complete the research phase and construct the final evaluation report.

The research design included fieldwork in four case study sites around the province. The findings were generated from fifty-four formal interviews from a strategic sample of actors, informal interviews and several meetings
in the case study areas. Also we used information located in CMHC and project files. In the final division of labour two members of the study team did the case studies outside the lower mainland, working together on both sites to shorten the time in the field. Two other members worked separately on two other sites within the lower mainland and I worked on the CMHC files and analysis of schedules as they were completed. None of this work was accomplished without a series of disagreements about who should be doing specific tasks and how those tasks should be accomplished. The process of negotiating became extremely complex and always included some persuading, a bit of educating, a lot of manipulation, and in some ways a kind of coercion on the part of certain members through the simple tactic of dragging on our meetings endlessly. This was the same member who seldom did his part of the work; he did not fulfill his interviewing assignments; he never completed his data gathering assignments; and he constantly provided confusing skimpy material for assignments that he volunteered to do as part of the text of the report.

Unfortunately there were no sanctions which could be applied internally to convince him of his responsibilities to the study team, nor was there any form of coercion that would have forced him to improve the quantity and/or quality of his work. It remained a "stand-off" and ultimately other members of the study team completed his interviews and rewrote his draft material. This collective inability to negotiate a compromise with a study team member seemed to reflect our inability to resolve other conflicts with the management team directing the evaluation study nationally.

Despite all these difficulties we continued to collect the interview data, financial information from project files and observations from meetings with CMHC staff and program participants in the case study areas. There were interviews with a wide range of individuals who played various roles in
the four case study areas: municipal politicians, municipal staff, NIP coordinators, RRAP coordinators, NIP committee members, other involved citizens (including bankers, realtors, media people, school officials), social workers, and/or community development workers. The specific nature of NIP in the four case studies determined which particular combination of actors was interviewed and how much time was spent at the case study site.

Organization of the Final Report and The Action Plan Meeting

After extensive discussions of alternative ways for analysing the interview results, we decided that the number of interviews was too small to make computer analysis appropriate. The process which we decided to use involved a general content analysis of the results in each schedule according to a series of topic area headings derived from the initial set of seven research questions. The objective data formats were not used, although a certain amount of that type of material was available from the various supporting documents from each case study area.

Because of the "richness" of the data obtained from each case study area we found that a new set of problems developed regarding the nature and structure of the final report format. We were given a great deal of material that was critical of the various actors involved in the NIP/RRAP programs as well as the functioning of the programs themselves. We had already guaranteed confidentiality to all the respondents with whom we talked, therefore, it did not seem feasible to return the interview schedules with these private remarks on them to the Head Office of CMHC. Instead, we wrote a regional report which set out the results in such a way that the necessary confidentiality could be guaranteed. We presented a great deal of detail where there was no problem of betraying anything that was said in confidence.
Our defence was that we were evaluating a program, not the specific projects. We wanted the report to be used as a general critique of the structure and functioning of that program.

Once the report was completed and sent to Ottawa in December 1976 we had one other chance to provide "input" into the continuing policy debate regarding the fate of NIP and RRAP programs. This chance developed when we were asked to send two representatives to an "Action Plan" meeting held in early March 1977 at Montebello, a rather posh conference centre located just outside Ottawa. When my partner returned from that conference she indicated that the NIP program had little chance of being extended past its planned termination in March 1978.

4.5 The NIP/RRAP Evaluation: A Study in Ambiguity and Antagonism

To the casual eye, it might appear that the Stage II SRP Evaluation and the British Columbia component of the NIP/RRAP Evaluation were quite similar. These were both assessments of programs concerned with housing and neighbourhood rehabilitation; they were both funded by the three levels of government and included some citizen participation; the evaluation studies included overlapping personnel and time-periods; and the research models entailed analysis of social, physical, cultural and economic aspects of the programs as implemented in the field. Yet, despite these overt similarities, there were several distinct differences in the negotiation context of each evaluation which resulted in significantly different patterns of cooperation and conflict among the main negotiators. Strauss's theories of negotiation context help illuminate these patterns in the NIP/RRAP evaluation.
Basic Areas of Negotiation

First, it was never clear just who all the negotiators were. Our main contact, the person who hired us, was Nils Larson, the research director of NRID. As well, there were the two members of CPAC, whom Larson, without consulting us, foisted upon us. The national evaluation was coordinated by an Ontario-based sociologist, who had the contract for developing the twenty-one case studies across Canada (in British Columbia, Alberta, Saskatchewan, Ontario, Quebec and one Maritime province). These were the main actors, the main negotiators. However, we did have quite a bit to do with Victor Virak who was head of the NIP program operations in British Columbia for CMHC and several other people at CMHC who were extremely helpful in getting the data base we needed for our case studies.

Not only were the negotiation sequential but they were also linked with negotiations at CMHC head office in Ottawa, negotiations about which we received little, if any, information. As well, a substantial amount of the activity was covert. Larson found it extremely difficult to make decisions; the national coordinator would frequently fly from Toronto to Ottawa, ostensibly to lobby for himself and his own particular approach; consequently we never knew where we stood. For example, one Friday afternoon in September we thought we had a contract for putting together the interview schedule for the RRAP evaluation; by Monday morning it became evident that we had no contract and that in fact the whole question of the RRAP evaluation had been dropped from the evaluation process.

In addition to the methodological problems discussed above, other problems with the national coordinator soured the negotiation process. We rarely had time to study adequately any of the provisions as they emerged. Quantities of new material -- all encased in yellow plastic three-ring binders -- were often presented to us on the same day as meetings were
scheduled for discussion of the material; most of it was exceedingly detailed and repetitive. We would have no time to review the material; we were also unclear where it fitted into this enormous intricately designed theoretical schema which appeared piece by piece; and were unaware of any kind of field testing that had been done with any of the complex interview schedule materials that were sent out. (At one point we were supposed to use sixteen different interview schedules to interview sixteen different kinds of "actors" in the NIP process. We managed to reduce that to one core interview schedule with some variations).

In retrospect it is evident that there were other problems which developed in the evaluation context. Some of the difficulties had to do with the covert nature of the negotiations that went on. Others related to the fact that the national coordinator was perceived as having extremely limited experience in field research in comparison with the members of the British Columbia study team. In addition, there was a lack of clarity and a lack of definition of the boundaries of who would be doing what during the course of the evaluation. And since there was no clear procedure for negotiating, making decisions, coming to some conclusion on any of these issues, we of the West Coast Mafia, using our label, found that we could not come to any kind of clear-cut negotiated conclusion with these kinds of issues. We therefore tried to overcome the problems on our own.

The national coordinator's "rational model" of organization management and control (Kanter 1975:43) tended to disrupt the development of cooperation within the study team. The implicit notion that the unequal distribution of authority aided the efficiency of the evaluation study worked counter to our generally non-dogmatic, non-hierarchical approach to accomplishing research. This preferred approach was documented in our experiences on the SRP evaluation. In the Strathcona team there had been nine main members
with differing disciplinary backgrounds but an underlying set of perspectives (political and/or personal) that made it possible to work together. All members held relatively flexible non-authoritarian theories of negotiation. That is, none of the people on the study team appeared to be so firmly attached to his/her point of view that he/she was not willing to negotiate, to consider some criticism, to consider an alternative perspective. There was a great deal of give and take on that study team and a substantial amount of back and forth learning. This could be seen in the remarks that people made in the work sessions and the backup that all the subconsultants in the team gave to Laurie Hurwitz and me in showing up at the final session and in defending the final form of the report. They all seemed to feel extremely positive about the work that they had put into the Strathcona evaluation.

In contrast, the study team in this NIP/RRAP evaluation became a most divisive group when it came to questions of the division of labour, interpretation of field data, and the construction of recommendations in the final report. One classic example of our dissension is that most of what appeared in the final draft of the evaluation report was determined by who had the toughest constitution and could manage to stay up the latest. As well, the balance of power seemed at times to be based on some apparent distinctions between male/female negotiating styles, the former being product-oriented and the latter tending to be a process orientation (Tresemer 1975).

Moreover, there were no decisions by consensus. In Strathcona, decision by consensus (no matter how minimal that might be at certain times) was the basis for decision-making in the SRC and, to a great extent, consensus was the basis for decision-making in the study team. In the NIP/RRAP evaluation, there was no consensus at the national level and there was very little consensus, comfortable consensus if you will, at the local level within the
team. While the team did band together in terms of finding strategies for overcoming our difficulties with the national coordinator there was always a high degree of interpersonal difficulty in producing those strategies.

These problems were made more acute by the general inaccessibility and frequent invisibility of both day-to-day and longer term transactions among the management group. We found it increasingly difficult to be overt or direct because of the covert "deals" which appeared to be constantly taking place in Ottawa. The distinctly covert character of much of the transactions disturbed certain members of the study team who had worked on the Strathcona project where so much of what happened was quite "upfront". In addition, the indecisiveness of Larson made everything much more complicated than it needed to be.

What was the nature of the respective stakes in the negotiation? Money was clearly an important stake: we had a contract to do an evaluation and we wanted to get paid for the work we were doing. We felt that the NIP program deserved a good evaluation in British Columbia, good in the sense that it deserved a careful critical evaluation that identified both the effective components and the difficulties. For Larson it appeared to be part of his job, and his problem was that he had to keep things moving along and reduce controversy. But he did not know quite how to deal with conflict. As for the national coordinator, he had been earning substantial consulting fees from CMHC over the preceding three years and presumably did not want that agreeable arrangement put in jeopardy.

With regard to the balance of power, it is significant that Ottawa began to refer to us as the West Coast Mafia. Generally we were not accepting what we perceived as essential inadequacies of design and implementation. Power was obviously located in the hands of several unseen negotiators at CMHC and finally at the Treasury Board. We had few options in
this situation and the small amount of power we had at the "bargaining table" was used to negotiate (or cease to negotiate) on three main issues. The first was, of course, the interview schedule which we completely rewrote and used in the four case study areas in British Columbia. The second instance was the RRAP impact study design, which we assessed with the help of two urban land use economists. We refused to implement that design in the field since it was shown to be enormously expensive and likely to produce relatively useless data. And, finally, when Ottawa insisted that we send along the interview schedules which contained raw data detailed enough to identify respondents, we refused to release the schedules. Instead we produced a comprehensive, detailed evaluation of those aspects of the program delivery system which were of greatest concern to CMHC in Ottawa.

In 1978, when the evaluation studies had been completed and the legislation came up for review, the senior executives of CMHC in coordination with several provincial Ministers of Housing decided that:

1. The NIP program be discontinued while reduced funding for these types of projects be given directly to the provinces in the form of "global agreements", a type of block grant.

2. The RRAP program be continued as a separate program under CMHC funding and municipal administration.

It has been hypothesized that these decisions were made on the basis of limited discussion and a series of rumours which said that the market for NIP had been "saturated" (private communication, CMHC Social Housing staff member). It would be difficult, at this point, to determine the full explanation. Perhaps something in the evaluation studies provided the provinces with additional material for their ongoing balancing act with the federal level of government. However, nothing in the brief contacts between the British Columbia study team and CMHC in Ottawa allowed us to follow up on the use/lack of use of our NIP/RRAP evaluation in the "corridors of power."
In this second case study, the evaluation of the NIP Delivery System in British Columbia, we find a significantly different pattern of negotiations from the SRP Stage II Evaluation. To begin with, the overall evaluation strategy for NIP had been developed by CMHC staff in Ottawa in conjunction with the national coordinator before any other areas were considered for negotiation. Authority and financial resources were centralized in the CMHC head office, and decisions were frequently announced without participation of any other level of government or community group. Funding was allocated to the British Columbia component of the national evaluation and split between CPAC (British Columbia) and Ptarmigan during a confusing hiring process that had no identifiable proposal call, interviews or other traditional elements.

The original arrangement presumed that control of research results lay totally in the hands of the NRID research director and the national coordinator, while the study teams in each province were completely accountable for bringing in the research data that was required by these two major negotiators. The presumed balance of control and accountability was directly challenged by the British Columbia study team in an ongoing series of confrontations. Based in two substantially different models of research, the conflict between the national director and the British Columbia study team moved from the conceptual to methodological areas and gradually into other aspects of the research process. The refusal to do the RRAP evaluation on the grounds that it was too high in costs and too low in benefits was one example. Finally, our decision to retain the original interviews and not send them to Ottawa (to protect respondents' anonymity) was the last in this series of objections to the controlling approach taken by the national coordinator.
Negotiating the overall division of labour in the NIP/RRAP evaluation was a complex and difficult process. In many instances the conflict was covert, since the coordinator tended to go directly to Ottawa to reinforce his own position. At one or two points members of the evaluation team in Vancouver threatened to quit. This option was never taken seriously but nevertheless it was one of the items put on the negotiating board. The fact that we began to work around the demands of the national coordinator symbolized how difficult it was for us to find any effective means to communicate with him directly. Furthermore, several schisms developed within the British Columbia research team which prevented any "web of commitments" (Strauss 1978:139) to grow up and be continuously renewed among the members. The unpredictableness of working relations with the evaluation management seemed to be an obstacle to achieving anything more than a "united front" when dealing with the opposition.

In summarizing the NIP/RRAP evaluation process I would argue that the effective terms include: ambiguity, conflict, manipulation, antagonism and confrontation wherever possible. This appeared to be an evaluation research setting where avoidance and rejection of any overt form of negotiation was the central characteristic of management personnel. The study managers tried to be oblivious to the British Columbia study team's claims regarding their own experience, competence and collective strength. The national coordinator in particular constantly strove to impose his vision of how the research "ought to be done" on a group of field researchers who claimed greater familiarity with the "policy territory" and more experience with the problems of empirical research in diverse urban and rural settings.

There are several generalizations which may be drawn from the evaluation research done in this case study. There were certain properties of the negotiation context which were evident in the very first set of transactions
and remained constant throughout the later stages of the evaluation. Specifically I refer to the ambiguity in defining the major actors and what I would call the "superrational" model of management and research design adhered to by the national coordinator of the overall evaluation study. These contrast with the large, but obviously identifiable cast of characters in the SRP evaluation and the relative research autonomy experienced by the study team working in that setting.

In addition, despite (or perhaps because of) the conflict between the national coordinator and the British Columbia study team regarding the appropriate research strategy (implicit in our attitudes towards respondents' competence and right to confidentiality), model and methods to be employed, the field research was completed with dispatch and the final report reached Ottawa just ahead of the deadline. This contrasts dramatically with the protracted negotiations of the SRP evaluation and the appearance of the final report nearly a year after the anticipated deadline.

Finally, with regard to both case studies, we must recognize that the evaluations were located at some distance from the policy intervention process and the evaluation reports were never utilized. This lack of research utilization must be taken into account when analyzing the overall structural context of policy development regarding housing problems in Canada. One may begin to question the evaluation process, its costs and consequences, when the products of evaluation studies are designed to be obsolete even before completion.

Before dismissing the potential of evaluation research as a tool in the policy intervention process, there are certain aspects of this NIP/RRAP evaluation case study which have negative but nevertheless important implications for the construction of additional models of knowledge production and use in policy research settings. The main question is whether it is
appropriate for policy research, and particularly program evaluations, to be centrally designed and administered even when the program to be evaluated is national in scope and funding. The problems of regional variation may be more suitably dealt with through a decentralized process for the development of research strategies, models and methods. It would appear that the control, objectivity, and generalizability which are supposed to be the products of uniform design and administration may perhaps be "false commodities". By that term I mean that such national evaluations can never produce a uniform, generally acceptable knowledge product, the cost and value of which is constant for policy-makers, program managers and program users all across the country. From the perspective of policy analysis such a product is impossible to generate. Therefore, another responsibility for policy researchers may be to educate their consumers regarding what their research dollars can buy and what they, as policy research funders, should be negotiating to receive as useful, well constructed research products.
1. Clairmont and Magill (1974) offer the most comprehensive Canadian analysis of the impact of urban renewal and the consequences of relocation in their study, Africville. Their work documents the life and death of the Black community of Halifax and details the evolution of Africville as a "social problem". The experiences of the Black community of Halifax were known to some of the younger, more political members of SPOTA (Strathcona Property Owners and Tenants Association) and provided further ammunition for their long fight against continued urban renewal (with relocation and redevelopment) in their own community.

2. I worked for the Vancouver Urban Renewal Study during 1968 as a consultant directing a survey of housing and neighbourhood preferences in the Cedar Cottage area. This put me in weekly contact with the members of the City Planning Department regarding policy issues and relations with the other two levels of government.

3. This is always evident in minutes and technical reports of Vancouver City Council as well as in interviews with several government and community representatives done during the SRP evaluation research.

4. The seminar held in May 1974 and reported by N. Bean (1975) could not be called an evaluation, but rather a "feedback" session for representatives of government from representatives of communities in which various types of rehabilitation had been undertaken with differing levels of grassroots participation. There was no systematic data collection or analysis involved in this effort in communication and information exchange.

5. The Stage I Evaluation of the SRC done by the then director of research for the United Way (using the resources of a management consulting firm) was late in arriving and relatively fragmented and superficial in its analysis of program operation and service delivery. Much of it was descriptive and pointed to those aspects of the project which were unqualified success. No overall analysis of the project spending was attempted and no contradictions or conflicts among the partners were brought to light. The Stage I Evaluation took consensus at face value and provided the SRC with an evaluation which did not attempt to deal with obstacles to participation or limits to project success.

6. Ptarmigan Planning Associates is a private consulting firm of which I was originally a partner, with Laurie Hurwitz an urban planner, and now am sole owner. At the time of our incorporation, November 1975, there was only one other planning consulting firm in Vancouver headed by women. I had done several years of consulting work on various OEO programs in the eastern United States, but only on short-term contracts for personal service. This was the first "entrepreneurial" experience for Laurie and myself. We had very limited assets.

7. He was replaced as chairman by Tony Geach when Pickstone left the City to go into private consulting. Geach became the Assistant Director or Planning and chairman of the SRC in June 1975 and was considered a much less competent and skillful negotiator than Pickstone, his predecessor.
8. Laurie had flown in from Bamfield on the west side of Vancouver Island by small plane and water taxi. She was taking a course in marine biology (an avocation) but the demands of Ptarmigan and the consulting business required that we both present ourselves at these negotiations. The three hour trip from a remote part of the Island to a small house in Strathcona required a rapid shift in mind set and appearance.

9. One of the less appealing aspects of Vancouver City Council and its committees is the extent to which they tend to "rubber-stamp" the recommendations and budgetary demands of the City Engineering Department. Following up and documenting that assertion would make a fine dissertation for someone else.

10. These budget analyses, done by two economists who were members of the Department of Commerce at UBC, and "experts" in urban land economics, were a very important surprise to the study team. They were our main "ammunition" in confronting the rumour circulated by several bureaucrats and West Side residents that "those old Chinese guys had ripped off the government again". This type of prejudiced remark had been freely offered by several civil servants indirectly connected to the SRP.

11. Curiously enough these notes appear to be on the back of some scrap paper that involved a very esoteric discussion given by somebody on the theories of Althusser and Bachelard, the French Marxist theoreticians.

12. We taped many of our work-group sessions. In listening to these meetings it becomes relatively clear that each member of the working group felt that she/he was contributing to a collective product. There was a great deal of discussion back and forth and some negotiation about how much time each person was able to put in on particular tasks. In general much of the activity worked fairly smoothly.


14. See CPAC (B.C.) and Ptarmigan Planning 1977 for details. An overview of the issues is included in the discussion of the methodological appendix to the study contained in this chapter.

15. Although the writer(s) of this handbook assure us that "the program is not designed to be used for wholesale demolition of deteriorated buildings and the subsequent construction of massive new building projects", the conceptual approach and lanaguage of the handbook retain the medical/organic model of growth and change in urban areas. The underlying image is the "sick body", and the "cure" which originally entailed massive surgery, now takes the more preventative, limited scale, and holistic approach of recent medical ideology.

16. By the time those meetings were completed we had begun to question the competence and analytical skills of the national coordinator. Members of the research team felt that he lacked the ability to communicate effectively in the research process: his manner was condescending, his theoretical statements were confusing and appeared unrelated to the original purposes of the study, and his specific research strategy was never explained.
16. His "theory of negotiation" was non-existent, since his approach was to impose the methodological framework and instruments without any opening for discussion. The results of this antagonism are most fully expressed in the section of this chapter dealing with the methodological appendix to the NIP evaluation study.

17. This material is drawn from CPAC and Ptarmigan 1977, Methodological Appendix. That appendix was written by me in consultation with the other four members of the study team. It was our attempt to document and make evident to the Ottawa CMHC staff the lack of competence of the national director and the inappropriateness of his research design.

18. Her main conclusions from that meeting were that: (a) there were significant differences between the urban and rural areas with regard to the operation of the NIP program; (b) when the goals of the NIP were in concert with the existing goals of other agencies then the NIP program worked, but it did not work under other circumstances; (c) participation varied extremely significantly from municipality to municipality; and that (d) some kind of basic community development work was essential to the effectiveness of the NIP program.

19. To some extent we were backed, though rather indirectly, by Lloyd Axworthy at the Urban Institute in Winnipeg who shared some of our difficulties with the research design that had been presented by the national coordinator.
CHAPTER 5: PROGRAM SCALE, LEVELS OF INVOLVEMENT AND THE SUBSEQUENT USES OF EVALUATION STUDIES IN ONGOING PROGRAMS

This second set of case studies will again be described and analyzed using the framework developed in Chapter 3 which focuses on the structural context of each program evaluation and presents the evaluation process in terms of a series of six areas of negotiation which may emerge as issues in the course of accomplishing the evaluation research and constructing and using the evaluation product. The case studies presented in this chapter differ significantly from those in the preceding chapter with regard to characteristics of program scale, substantive policy area, funding arrangements, time-frame, and the size of the evaluation study team. Whereas the first set of case studies involved relatively large-scale programs, reflected shifts in national housing policy and entailed cost-sharing by all three levels of government, the studies presented in this chapter involved programs that were smaller in scope, concerned social service and educational responsibilities of the provincial government, and entailed cost-sharing among provincial, municipal and local community funding sources.

Specifically, the evaluation studies described in Chapter 4 concerned housing and environmental change, while these case studies focused on programs concerned with the delivery of social services in an ethnic community and the provision of support services in legal education to all secondary schools in British Columbia. As well, in the case of the Strathcona and NIP/RRAP evaluations, the initiation and funding of these studies came from the federal government bodies which had prime jurisdiction over the programs. In the case of the evaluations discussed in this chapter, the request for evaluation came from within the programs themselves; funding came from the operating budgets of the two programs. Moreover, at the time of the evaluations of the SRP and NIP/RRAP, the program operation had either been completed (SRP) or was coming
to a close (NIP/RRAP). In contrast, in this second set of cases, the evaluations were commissioned for ongoing programs already in operation; they reflected an attempt by those involved in the programs to develop new policy and direction for those programs. One final contrast separates the two sets of evaluations: in those described in Chapter 4, the evaluation teams were large. Nine were involved in the SRP evaluation, while five plus support staff worked on the B.C. component of the NIP/RRAP evaluation. In contrast, in the first of the case studies described here, two of us did the evaluation; in the second case, I alone was involved and then not as a formal evaluator but simply as a 'consultant'. In the former cases, the research staffs had no further input into the policy intervention process after their reports were produced; in the latter, the program evaluator(s) continued to be consulted in both formal and informal ways as new policies were developed and organizational changes were funded and implemented.

5.1 Case: An Assessment of the Vancouver Jewish Family Service Agency

As background to this particular case study I want to present a brief historical description of the philanthropic tradition among the Jews of Eastern Europe before the revolutions and enlightenment of the mid-nineteenth century. Moses Rischin (1964:37) offers the following summary of the value system and structure of philanthropy which was the "ideal" of the period before the Eastern European Jews became integrated into the class struggles and cultural transformations occurring in Russia, Poland and Germany throughout the last half of the century:
Beyond the daily devotionals, religion bound its ethical prescriptions into the social fabric, making charity an obligation. Voluntary societies, or hevras, flourishing in every hamlet and shtetl (town), competed in raising free loans for the needy and dowries for the orphaned girls. These hevras aided the sick, boarded the transient, and educated the children of the poor. Honor and social prestige were the reward for such good works, to say nothing of benefits to be derived in the world beyond. Although the merchant of means was set off from the artisan, the townsman from the villager, and above all, the learned from the unlearned, neither erudition nor wealth conferred caste privileges. The blessings of affluence, though hardly undervalued in worldly terms, had a social and religious purpose, that of tsedaka, charity conceived as justice. Tsedaka was to be acquitted with infinite tact, personal generosity, and humane feeling for the less fortunate of Israel's children. This social responsibility, which prosperous Jews were to bear with humility and self-effacement, served as the standard of the world's respect. If practice deviated from precept, if wealth at times showed itself neither generous nor kind and learning neither wise nor virtuous, still these ideals established the moral tone. Even the most delinquent and least literate of Jews was measured and measured himself by the ideals of Torah.

The Structural Context: Public/Private Social Service, Community Organization, and Patterns of Charitable Contributions

In British Columbia the State delivers the basic social services (e.g., welfare, aid to the handicapped) through the Ministry of Human Resources (MHR). The Ministry also pays for other types of social service on a yearly contractual basis (e.g., daycare, Family Centres). In addition, there are a wide variety of private service agencies which receive no substantial government support but have to rely on the yearly fund-raising drive of the United Community Services (UCS) and other charitable donations. The Jewish Family Service Agency (JFSA) of Vancouver is one of the latter category of private social service agencies.

It is one of several private ethnically or religiously assisted agencies whose activities are intended to supplement or expand upon the basic level of service provided by MHR. In 1977, when the evaluation research was done, the JFSA had a total budget of approximately $55,000, of which 60 percent came from UCS, 18 percent from a special endowment account, 12 percent from
charitable donations given privately throughout the year, and 10 percent from the Jewish Community Fund and Council (JCFC). The latter is an elected board which administers all the donations from United Jewish Appeal that are earmarked for local community use. Thus JCFC allocates funds for a variety of communal institutions including: JFSA, the Jewish Community Centre, the Home for the Aged, the Talmud Torah (a religious school), and a variety of other related social services (including two Jewish camps for children and Kosher Meals on Wheels).

Certain characteristics of the structural context of the Jewish community in Vancouver relate to the whole question of fund-raising for social and educational services. The Jewish community in Vancouver is renowned for raising substantial amounts of money to be sent to Israel, but it gives just about the lowest rate of support to local services and facilities. Whereas in typical North American Jewish communities 25-30 percent of the private funds that are raised are kept for local needs, in Vancouver the percentage kept for local needs is somewhere between 15-17 percent. That figure cannot be documented totally, but from interviews with local fund raisers this was the estimated figure.

In recent years, the community has expanded in size and diversified in terms of the nature and ethnic backgrounds of the Jews that now live in the Lower Mainland. There are differences in their family patterns, in their class backgrounds, and in an assortment of attributes that one could call cultural characteristics. In general up to fifteen years ago the community was relatively homogeneous, dominated by a group of upper middle-class families who had made their money in other parts of Canada and then moved to Vancouver. Most Vancouver Jews were from Eastern European backgrounds originally and were first or second generation, but a substantial proportion of them came
to the city, having assimilated to the dominant Anglo culture and achieved economic success in Montreal or Winnipeg.

But the situation has changed. In the last fifteen years there has been an influx of new varieties of Jews, from Israel, Russia, the United States, England, South Africa and South American countries. While they represent a wide variety of class backgrounds, there are clearly many more working class Jews among them. There are many more Jews with trades and manual skills, whose backgrounds are not in what has been traditionally stereotyped as Jewish occupations. For many of them English is not the first language; for the North African Jews, for example, Arabic is the first language, and for others Russian or Yiddish are the first languages. In addition, over the last fifteen years the increasing rates of intermarriage and divorce have significantly affected the nature of the local Jewish community (S. Landau and S. Barnett, personal communication). All of this has created a new heterogeneity in the community. For example, there are now larger numbers of non-affiliated Jews; that is, Jews who identify themselves as Jews but are not affiliated with any of the synagogues or temples.

In many ways the community is very different than it was fifteen or twenty years ago. However, that change has not been totally reflected in representation on the JCFC Board, especially up to the late 1970's. The lack of recognition of new social needs had some long-term consequences both in terms of the JFSA, and in terms of the delivery of other kinds of cultural/educational services within the community. Several community members trained in social work, as well as other social sciences, argued strongly for a shift in the allocation of community funds so that local needs (cultural, educational, training and social) could be more satisfactorily supported. But the power elite within the community (male, middle-aged, and economically upper class) could not be convinced of the need for expanded services and support systems for newer
members of the community, for the steadily increasing numbers of elderly Jews requiring health care and emotional attention or for single parent families.

Preliminary Negotiations: The Outside Evaluators as Determined "Outsiders"

In early 1976, the Jewish Family Services Agency faced a dilemma common to many social service agencies: What happens to such an agency when it is continually told that it serves few community needs but is at the same time underfunded and understaffed to meet the needs of its existing clientele?

Answer: It has itself evaluated and uses the resulting document as a basic piece of ammunition in its fight for survival and perhaps even expansion.

While an oversimplification of the structural and negotiation contexts in which the JFSA assessment was done, it does suggest the situation in which its board decided that only a document produced by "independent evaluators" would serve to legitimize their continued requests for grants and charitable donations.

The JFSA initiated indirect contact with Ptarmigan Planning in July 1976 when Martin Zlotnik, JFSA board president, contacted Michael Goldberg, an associate who had worked closely with us on the preceding evaluation studies. Initially he and Zlotnik discussed the desirability of an overall needs assessment of the three Jewish social service agencies in Vancouver, namely, the JFSA, the Jewish Community Centre, and the Home for the Aged. But lack of support from the Jewish Community Fund and Council Board for this broad evaluation, which we estimated to cost between $14,000 and 18,000, narrowed the assessment to one of the family service agencies alone.

The terms of the research were clear cut and specific, focusing on the following questions: 1(a) What are the stated goals and objectives of the JFSA?; (b) What services and programs are provided to meet those goals and objectives?; (c) What specific needs are being addressed by those services
and programs?; and (d) How do these compare to other JFSAs in Canada and in the United States? 2. Who are the current and immediate past users of the JFSA?; and 3. What are the perceived needs of people in the Jewish community and what is the perceived role of the JFSA in meeting these needs.

We clearly stated that the evaluation would "not include an assessment of staff performance, nor will it include an evaluation of the effectiveness of current service delivery". There existed at the time a serious rift between the Agency board and its director. Because the board lacked a personnel policy, it had no way of firing the director. Moreover, although board members were discontented with her approach to running the agency, they were unclear as to what kind of replacement they wanted. Our intention was to stay clear of this personnel conflict and not be used to do the board's "hatchet work".

Michael Goldberg and I were to do the work: his background included a B.A. in economics and an M.S.W. He had worked on the Strathcona study, had done a social needs study for the Strathcona area, and had worked on the NIP/RRAP evaluation. His training thus complemented my background in sociology and urban planning. As well, we were both Jews, educated in the United States. Our conservative budget for the proposed evaluation was $3,000.

However, it took more than a year of occasional letters and phone calls between the Agency board and Michael Goldberg before the President received sufficient support to proceed with the evaluation contract. During that time board members appeared to be totally involved in the controversy with their director and found it extremely difficult to make decisions of any kind. Finally, in August 1977, we began the evaluation research and completed the final draft of the report in December of that year.

The initial group of actor's represented in the evaluation process was very limited in number and diversity of interest. On the one hand there was the Board of Directors of the agency itself, headed at that point by Martin...
Zlotnik (who subsequently went on to be a TEAM candidate for Mayor in the City), and a group of about twenty members of whom only five or six were really active. Although composed of equal numbers of men and women with professional training (e.g., social workers, lawyers, business people, teachers) the board was considered relatively "low status" in the community, in that it did not have much political clout and it did not have any members of the financial elite among its members. There was no "consumer representation" on the board; in other words, nobody who used the services of the agency had been asked to join the board. Finally, members tended to be older people, with the youngest member being in his late thirties. The composition and status of the board was something that we were aware of and tried to deal with indirectly in the framework of our evaluation.

In addition there was the agency director and members of her staff. The director was in conflict with the members of the board regarding the scope of her control and responsibility in day-to-day operations. We did not want to know about the controversy in detail because we really did not want to get drawn into the personnel issue. There was a very small staff, a bookkeeper and a secretary and a few people around who did some volunteer work for the agency.

As the evaluation process began these appeared to be the only actors who had any interest in our process and results.

The Research Process: Methodology, Short-Term Consequences and Problems of Distribution

A minimal amount of negotiation centred around the basic research issues in the JFSA evaluation. Michael and I worked together and developed a conceptual framework and a set of methodological strategies which, by and large, were accepted in their complete form by the board and by the agency director.
and other people with whom we worked in the research setting. The underlying model was a fairly eclectic one which was basically a combination of a needs assessment and development of a series of alternative scenarios for the agency to move their activities through the next five years. The research design was relatively uncomplicated; using very straightforward methodological strategies, it was meant to collect information in specific areas.

The board essentially was interested in acquiring information which would assist them in their future planning for the agency. They recognized the need for a planning study because of the change in clientele being served by the agency. In the early 1970's, the agency had begun to serve increasing numbers of Jewish Russian immigrants, but by the middle of 1977 the flow of immigrants had been reduced and the board began to question their future direction. Under the terms of agreement with the agency board outlined above, we were to prepare a profile of current and past users of the JFSA; to obtain from interviews with "key informants" their perspective on the needs/problems of the Jewish community and their perspective on the role the JFSA could play in the community; and to compare the Vancouver JFSA with other JFSAs in Canada and the United States.

Since the methodology depended upon information from three sources we developed two questionnaires and an interview schedule: one questionnaire went to other agencies to obtain relevant information, one questionnaire was used to collect information from files, and the interview schedule was used with a series of key informants. We assembled the report using statistical frequencies and percents for certain kinds of information and content analysis for certain other kinds of information. We developed several interesting ways of ranking the types of problems or needs that were indicated. For example, in order to summarize our findings regarding the types of programs/services offered by the Vancouver JFSA and compare it with other agencies surveyed, we
developed a single table which ranked type of services according to the frequency with which they were offered by all agencies (core, secondary, tertiary, other) and then indicated which of these services the Vancouver JFSA was able to offer (Goldberg and Levitan 1977:36). The table underlined the absence of services to children and youth in the community (both in and out of school) and the lack of any form of homemaker service to the sick and/or elderly.

In addition, we developed a table which listed agency staff and budget by size of Jewish population for a series of nine cities including Vancouver. This was done in order to emphasize the lack of support in Vancouver and the really inadequate funding level of the local agency. Among the conclusions that we drew from that table (Goldberg and Levitan 1977:35): "There are wide variations in per capita figures for staffing and funding, but in both situations Vancouver is in the lowest category, i.e., less than 1 staff person per 4,000 Jews and expenditure of less than $5.00 per person per year."

Since the research design involved collecting data that would be aggregated, descriptive data on past and the present case loads, we made arrangements to have the members of the limited agency staff do the one sheet collection of information on each case. We felt that we could not go into the files, both in terms of the cost of our doing that work, and secondly in terms of maintaining the anonymity and the privacy of people who had cases with the agency. So after doing a random selection of numbered files, members of the staff were involved in collecting that data.

The director and members of the staff were quite open with us and we with them because we made it very clear to the board and to them that we would not intervene in whatever personnel squabbles they were having. There were repeated day-to-day activities with the director of the agency and staff in terms of data collection and explanation of procedures. There was very little in the way of negotiations with the members of the board until the report was
completed. We were given a great deal of freedom in terms of the data collection and we received a great deal of cooperation in the focused interviews that we did with a cross section of key informants in the community.

We gained a great deal of information about the image of the JFSA in the community, the needs in the community, the target groups that needed to be reached and quite a bit of unsolicited information that involved a critique of what was happening in terms of local services available in the community. Some of this information/opinion we were able to use in the evaluation report itself, and some we simply could not include because it went into enormous detail regarding some specific incident in the history of the agency.

The final report was sent with a covering letter to Marion Dewitt, the new President of the Board of Directors, on 25 January 1978. The letter pointed out that: "as indicated in the final chapter of the report, we consider that there are three main options available to the board of the agency. It seems evident to us that Option 3, an expansion of service, is the most desirable course to take, however, the choice ultimately rests with you and your board." Following that letter I have a response on 20 February 1978 from Marion Dewitt and she says:

I am pleased that you will be able to attend our meeting on 2 March. In the meantime you will be interested to know that the meeting of the representatives of the budget and allocations committee was held with JFSA and while we only used excerpts from your report, it was felt that the need for JFSA in the community had been 'professionally' established, and we have been asked to prepare the board that we will require further funding in order to put the most urgent needs as set out in your study into effect.

Later that spring Michael and I were invited to their annual general meeting to comment on the findings and options presented in the report. We were well received and there was little obvious controversy.

For reasons of costs, the report did not receive wide distribution. We had originally arranged to give the agency a basic copy; they would then
reproduce the desired number of copies at their expense. The report was reproduced but apparently there were few copies available, the reason was given as cost; although the report is not exceedingly long—it is 47 basic pages and some 10 pages of appendices. I was therefore personally approached by people for copies of the report, and I could not give them copies although I did find this peculiar. The one person that I did give a copy to was the Rabbi who is the Director of Hillel, who simply had been unable through the usual channels to get a copy of the assessment.

The above discussion makes clear that the JFSA evaluation posed few problems: there were no serious difficulties of a methodological kind, nor did Michael and I have any problems assembling the report and getting it to the board. There were problems of publication, as indicated above. Finally, the publication was used in program and policy change at the level of the agency itself, but it had virtually no impact with regard to the social structure of funding and policymaking in the larger Jewish community. The JFSA got a bigger share of what is still much too small an allocation of community funds.

5.2 The JFSA Assessment: Levels of Involvement, Evaluation Use, and the Legacy of an Unchanged Structural Context

In analyzing this case study there were several areas of negotiation in the evaluation process which emerged as issues in the short run such as the varying levels of involvement in the research process and the "other" negotiations which were going on between the board and the director. In addition, there was the more substantial problem of the structure of philanthropic resource allocation within the Jewish community which set up obstacles to any permanent improvement in the services offered by education/culture/welfare organizations oriented to local needs.
Basic Areas of Negotiation

The stakes of the evaluation team were fairly straightforward. Ptarmigan Planning wanted to produce an accurate, well-documented assessment of what the agency had been doing. And, second, we considered it essential to suggest some options for the board, director and staff to discuss and use as the basis for some long-term planning and fund raising. We already knew enough about the community from personal and professional sources to have a sense of the growing need for service. The JFSA assessment offered an opportunity to test those 'hunches' through a systematic analysis of the data we were able to collect from our sources. Michael Goldberg and I came to the evaluation negotiations with a sociologically sophisticated analysis of the workings of the community, as well as an inside knowledge that came from having grown up in families that had been the recipients of the services of JFSAs in other cities. We were not, then, typical upper middle class evaluators, although we both brought professional training and practical experience to the work.

As program evaluators hired on contract by the JFSA board we considered our evaluation work as essentially external to the organization, despite efforts to make us focus on problems of the internal division of control and responsibility. Our backgrounds and research experience sensitized us to the limits of employing a contentious research approach in this setting where a serious organizational conflict already existed. Furthermore, the lack of any basic monitoring procedures or methods to keep track of caseloads, services, and unmet needs suggested that what was most required by the agency was a review of their operations which was located within their existing conceptual/political framework.

In the initial phases of the evaluation research there were few other actors whose interests were directly represented in the negotiations. A few members of the board, on certain occasions the agency director and members of
the agency staff were involved in limited aspects of the research phase and their concern with the evaluation process varied considerably. There were differences of involvement among even those members of the board who had been relatively active in the preceding few years. Their level of involvement in the evaluation process varied across a range from "that's nice" to a core group of five members (those sharing the responsibility for decision-making) who participated in some areas of the design of the study and short-term interest in implementing the recommendations. One key member of the board took "ownership" and responsibility for doing the long-term planning, lobbying and negotiating to put the recommendations into practice.

The director had offered to assist us with the data collection process and spent a certain amount of her time organizing files and staff time. She also permitted us to interview her, and was quite candid about her long-term difficulties with the board regarding the scope and boundaries of her decision-making powers. She appeared to have or want very little stake in the evaluation process. Her level of involvement was a kind of passive participation, since much of her negotiating resources were invested in the conflict with the board.

Staff collected some data for us in the sample from the closed and continuing files. We also interviewed them and found that they were pleased and hopeful about the evaluation process. But essentially they defined the evaluation report as the property of the board. The management controversy and the hierarchical nature of the authority structure of the agency seemed to contribute to their positive, but essentially restrained approach to the evaluation process.

The interviews with "key informants" brought a new set of actors into the evaluation process. These community members, selected on the basis of their interest or involvement in related cultural and social services, brought forward their own substantive concerns with the way the community functioned and
the image of the JFSA. In addition they expressed their own preferences as to how our final report ought to be structured and which findings we ought to emphasize. Their preferences did help to guide our efforts in constructing the final report as a document that could be easily read and directly understood by a wide variety of community members.

In the course of our research it became evident that the negotiations regarding the agency assessment were only one of two distinct sets of negotiations going at that time. For several months before we received the request to go ahead with the evaluation proposal, the board and the executive director had been engaged in negotiations regarding the scope and limits of her power in administering the agency. These negotiations, while relatively covert, were ongoing, and serious for all parties concerned. The director appeared to be challenging the balance of power and demanding that the board develop a clear policy regarding her responsibilities and theirs. The board, which really meant only four or five active members, had no policy but appeared to dislike the director and considered her unfit for her job.

This critical problem within the agency faced us at the beginning of the evaluation process and continued for nearly a year after the assessment report was completed. The conflict between the board and the director was never resolved: a few months after completion of the evaluation, the executive director was fired; she in turn brought a countersuit for wrongful dismissal. Although the two parties reached an out-of-court settlement, the bitterness of this conflict paralyzed the board for nearly two years. Moreover, the basic issues of that conflict remained, since the board still has no effective policy for planning, implementation, and evaluation of personnel and programs.

The negotiations regarding the evaluation of the agency had a different focus and set of issues. We refused to evaluate the personnel situation since we defined it as outside the boundaries of our authority. We stressed that the
internal relationships of the agency were issues beyond our influence or control. Our unit of analysis was the agency as a whole, as it operated in the larger community context and in the context of other JFSA's in cities with similar size Jewish population. Obviously, there was a great deal of direct and indirect pressure to include the internal relationships of the agency as part of that unit of analysis, but we successfully resisted these pressures.

Both of us were firmly convinced that we needed to make our work as visible as possible, in order to defend ourselves against the threats and conflict between the board and the director. We did not want to be accused of doing any behind-the-scenes manipulating and always disassociated ourselves verbally from that controversy. Therefore we made it a practice to be very straightforward in identifying who we were, what we were doing, who had hired us, why we were doing it, and how the material would be used. We made it particularly clear that the materials collected from case files and interviews would be aggregated, and stressed that no individual would be identified by name in the final document.

In retrospect, it may be said that both sets of negotiations, the conflict between the board and the director, and the agency assessment process, have never been concluded since there is still no management policy or long-term planning process. Some or all of this legacy may be the result of several factors in the structural context of the agency which could never be remedied by a single assessment study. These factors include: the low level of financial and political support for local needs among powerful members of the community; the low levels of involvement in policymaking regarding local needs; the limited power and low status of the JFSA board members in the Jewish community; and the basic inability of these board members to define their own power and allocate clear responsibilities to their director and staff.
These factors, going from the general to the specific, were not fully evident four years ago when the assessment was completed. But the basic elements were there. And it is only by returning at this later date, through reading and interviews, that the overall pattern of connections between the structural and negotiation contexts of the original assessment became more obvious.

**Ongoing Problems and the Use/Misuse of the Report**

There were some specific short-term consequences of the evaluation study. First of all, as is evident from the board president's letter 20 February 1978 (see pg. 11), it was used in supporting successfully a bid for expansions of the JFSA funding base. In the year following the evaluation, the contribution from the Jewish Fund and Council was doubled and in subsequent years it was tripled. From the perspective of a program evaluator the evaluation and the options that we offered seemed useful in obtaining increased levels of funding.

There were as well, several long-term consequences that deserve mention. While the conflict between the director and the board had preceded our hiring as evaluators, a few members of the community (supporters of the director) were personally critical of us. They considered that the assessment had been used as the basis for her dismissal and therefore wanted to cast doubt on the legitimacy of our data and conclusions. As part of this process, the report was singled out by one particular academic as being invalid because we were "communists who had no feelings for Israel." Fortunately, this criticism seemed to have minimal significance for a great number of people concerned with services in the Jewish community.

The basic problems faced by the JFSA have, in the long term, remained. While the contribution of the Jewish Community Fund and Council increased from $7,000 to $48,000 between 1978 and 1980 and the agency also received a
substantial increase from the United Community Services, this was achieved only as a result of strenuous lobbying. The person responsible for this success was Shirley Barnett who became JFSA board president and acting executive director after the out-of-court settlement with the previous director. Her leadership abilities and competence appear to have been the key to the short-term improvements in funding, service and direction. However, as she was quick to point out in an interview I held with her in late 1981, these abilities produced their own difficulties within the negotiating context of the JFSA. While other members of the board and staff constantly relied on her, thus increasing her power, they resented her power. When an outside consultant, an expert in management/staff relations in service agencies, was brought in to try to resolve some of these difficulties, so much hostility was directed at her that she immediately withdrew from involvement as president and acting director. Another period of immobilization set in before the board was able to decide to promote one of the part-time social workers to be executive director.

Despite this change, however, the problems remain. Conflict with other beneficiaries of the Jewish Community Fund and Council is ongoing, the other agencies resenting the increased funding given to the JFSA. As indicated by several of our "key informants" that conflict is rooted in the basic inequities of the funding policies of the Jewish community; that is, the limited percentage of the overall charitable donations that is retained for local use.

According to the present executive director, Susan Landau, whom I also interviewed in late 1981, the evaluation itself has caused a problem within the agency. Four years after its completion, it is viewed by several members of the board as the last word on what the agency should be doing, whereas she recognizes the need for ongoing planning and implementation and evaluation of services. In her opinion, an evaluation that is four years old is not particularly useful if one is going to have organizational change and appropriate planning for service delivery.
She also emphasized that in general, community funding priorities have remained the same as four years ago, that is support for Israel with inadequate support for local needs. Her warning is that in ten years support for Israel will be gone from the community if there are no adequate resources for maintaining Jewish culture here in Vancouver. Furthermore, there has been no significant change in the negotiation context of funding for local needs since the members of the elite who control the Jewish Community Fund and Council remain the same or come from the same economic backgrounds. The newer ethnic members of the community, ethnic in the sense that they come from other countries, and the newer elements who come from working class backgrounds, have begun to increase in number and have begun to solidify. Therefore, it becomes even more apparent that their needs, their cultural, social, psychological, employment needs are simply not being met by the services available in the community.

From her perspective as the present JFSA executive director, it is apparent that the evaluation met the short-term needs of increasing the funding available to serve immigrants, particularly Russian immigrants, and to serve more seniors. And, in looking over the needs and services profile developed in the summer after the evaluation was done, it is clear that the increased staff level is able to meet some of these needs more effectively. However, in the long term, the interpretations and options presented in the 1977 assessment have become enshrined as "the truth" for certain members of the board. Thus it has become an obstacle to continuing a process of ongoing planning, implementation and evaluation. Consequently, management difficulties and funding problems, both of which seem to derive from the ideological position taken by members of the power structure of the Jewish community, still remain.
Summary and Implications for Knowledge Production and Use

This case study documents the evaluation of a private social service agency using a research approach that took existing organizational goals and operations as "givens", used a methodology which utilized existing records and record-keeping methods augmented by interviews with fifty key informants and comparative data from other similar agencies across Canada and down the west coast of the United States. The resulting evaluation document was used effectively in gaining increased funding for the agency. However, the document had no long-term effect on the unclear division of power and responsibility between the agency board and director. Furthermore, criticisms included in the research report apparently had little or no effect on the structural context of community resource allocation.

There are two characteristics of the structural context of the JFSA, and the overall research design which may be considered as contributing to the overall pattern of short-term impact but the minimal long-term effect of the agency assessment. The first characteristic of significance was the differential levels of involvement in the JFSA evaluation process and the second was the unit of analysis, a single agency, which became the focus of what should have been a more comprehensive process of community needs assessment.

When we examined the levels of involvement of the negotiators in the JFSA evaluation, we found that these differed markedly among individuals at varying levels in the organization (staff, director, board members) and with regard to particular areas of negotiation (hiring, research design, implementation of recommended changes) in the evaluation process. For example, all board members were consulted regarding the use of Agency resources to commission an external evaluation. However, only a selection of members continued to participate through the hiring process and went on to approve the research design, were
interested in the study results, and stayed involved in the process of setting new priorities and encouraging Agency growth and development.

As indicated in 5.2, one board member decided to take ownership of the final report and spent the following two years spearheading a push for increased funding and expansion of Agency staffing and programs. There was, as mentioned, a kind of "backlash" among other board members regarding her seizure of power and authority. However, it is important to understand the relatively low level of participation that characterized most board members over long periods of time. It would seem that the very high level of involvement of one member tended to accentuate the relative passivity of the others. This awareness of an imbalance in the distribution of power and participation was brought out and accentuated in a workshop designed to examine board interaction and process considerations. The negative feedback to the highly involved member of the board resulted in an immediate withdrawal of her energy from Agency responsibilities.

Staff members and the director were consulted at the beginning of the evaluation process and the board delegated to them certain responsibilities in the data collection process which resulted in a form of participation. But their level of involvement was relatively low in all the other areas of negotiation in the evaluation process. Of course, the director was much more fully involved in her controversy with the board regarding the scope and limits of control related to the running of the Agency and negotiations concerning her job tenure. This was a serious obstacle to her involvement in the evaluation and was a major problem around which staff members had to negotiate their day-to-day activities.

In examining the overall pattern of involvement in the JFSA evaluation, we can recognize that the marked variations in involvement tended to come from two sources. First, although the Agency is quite small in size, it is characterized
by a hierarchical social order that has been continually disrupted by an ambiguous distribution of power and responsibility. This problem has tended to distract members at all levels from fully involving themselves in the long-term tasks of Agency planning and program change. Second, the relatively low status of the Agency within the Jewish community has tended to contribute to the low level of involvement of board members in the responsibilities of making and implementing policy for Agency operation. Both these factors in the structural context of the Agency contributed to a general lack of involvement in the development and subsequent use of the findings in the agency assessment.

Clearly, the evaluation of a single agency was not sufficient to make an impact on the financial elite of the community. Our original discussion of an overall needs assessment of the three Jewish social service agencies in the city (the JFSA, Jewish Community Centre, and Home for the Aged) might have had a greater impact since it would have involved a wider cross-section of the community and received greater publicity and distribution of research findings. In retrospect it becomes evident that the unit of analysis, a single agency, was inappropriate in scale for achieving any significant effect on those Jewish community policymakers who control resource allocation.

This question of the appropriate unit of analysis and the problem of defining the necessary organizational boundaries for evaluation research may become central issues in early planning for evaluation research use in the policy process. When studies are national in scope and employ a single research model there is the danger of overgeneralization and a basic lack of accuracy and validity of findings. On the other hand, where the study unit is limited in size and scope there is the objection that research results are idiosyncratic and do not permit generalization to the larger universe of similar organizations/communities as a whole.
5.3 Case: Legal Services Society Schools Program Evaluation Report 1980

The Structural Context: Stability and Shifts

Several educational and legal institutions as well as community based legal services and educational services in British Columbia do various forms of legal education. Some of them, specifically the law schools, do professional education. In the Faculties of Education at the University British Columbia, the University of Victoria, and Simon Fraser University, courses in legal education are available to those who will be teaching Law 11 or General Business 12 in the school system. As well, organizations such as the People's Law School offer specific short courses to the public while various community groups like West Coast Environmental Law educate both professionals and the public on questions which involve issues of public interest.

The Legal Services Commission was brought into existence by provincial legislation to provide legal information services to the public through a series of community law offices, a Native Courtworkers Program, an extensive legal resource library, and the Schools Program (See Figure 2). The mandate of the Schools Program was to "expand the quantity and quality of legal education in the public schools in order to facilitate pupil growth, development and understanding". By its third year of operation, the Schools Program included several components. It provided services such as courtwatching and teacher education programs; produced materials such as a newsletter and curriculum materials; and did promotion, sales and distribution for these services and products. In addition, it was connected to an internal (the LSC support system) and external (for example, the Ministry of Education, Bar Association, B.C.C.L.A.) liaison network of organizations which were all involved in some aspect of the field of legal education.
Because it is a legal service, the Schools Program comes under the jurisdiction of the provincial Attorney General; it is thus the only education program that remains outside the direct jurisdiction of the Ministry of Education. Although there is constant liaison between the director of the program and Ministry of Education representatives, there is occasionally pressure put on the project to have it relocate in the Ministry of Education. Such pressure has been successfully resisted; evidence from the evaluation suggested that it would become quite lost in the Ministry and that its present location is the most appropriate.

In 1980, provincial legislation amalgamated the Legal Services Commission with the Legal Aid Society to form the Legal Services Society of British Columbia. Peat Marwick Consultants Ltd. developed the organizational design for the amalgamation. It consisted of two components: 1) a revised management plan which merged the two original organizations only at the top with an additional layer of supervisory personnel (see Figure 3); and 2) the floor plan for a newly renovated and much larger set of offices and space for the resource library in a centrally located building in downtown Vancouver.

Both design components resulted in a variety of unanticipated problems for the new organization, ranging from the failure to include an effective system for taking and distributing telephone messages to a series of health problems related to the chemicals, lighting and air-conditioning system in the renovated office space. As well, when they all moved into the newly renovated offices, the staff of the formerly separate organizations had little contact with one another, except at the highest levels of management; they felt like the partners in a "shot-gun" marriage (Levitan, Field Notes, January 1980). Further, the Board of Directors of the new Legal Services Society included only a few of the original board members from the Legal Services Commission; the new
members, both from the Legal Aid Society and the Bar Society of British Columbia, had little or no knowledge or interest in legal information services. This put the director of the Legal Information Services (LIS) component under great pressure to explain and document the market for the services and to justify their significance to members who would have preferred to have the entire Legal Services budget directed into Legal Aid.
FIGURE 2 ORGANIZATION CHART OF THE LEGAL SERVICES COMMISSION, 1978-79

Funding Relationship

FIGURE 3 ORGANIZATION CHART OF THE LEGAL SERVICES SOCIETY, 1980

The Schools Program Evaluation: A Brief Chronology

The decision to have an evaluation came from within the organization itself. Since staff members had little knowledge of such a process, they invited me to meet with them to discuss whether and what kind of evaluation could be done to help them decide where the project should be going over the next few years. At this meeting which occurred in October 1978, I suggested questions to them (e.g., about the kinds of statistics they kept; about whether they had surveyed the teachers and students or parents who used their materials; and about their general overall relationship with members of the business, legal and educational communities) which would help them define their goals for the proposed evaluation. After two additional meetings in December 1978, I began the evaluation process with a contract for a certain number of hours of work during the first three months of 1979. My involvement with the project, however, continued throughout 1979 and into 1980; I reviewed the final draft of the evaluation report on August 12, 1980.6

Defining the Process: The Decision, Type of Evaluation and Consultant's Role

The desire of the Schools Project staff to do an evaluation stemmed from their increasing workload. Project staff were overloaded; that is, each day was full or "urgent" things to be done. All felt trapped by the immediate demands of the day-to-day operation. There was little time or energy to make decisions in terms of long-term goals, and no integrated perspective from which to say "yes" or "no" to these demands. They realized that the program had grown immensely in the three years since its inception: a variety of materials had been produced in various formats, workshops had been held, a newsletter developed, the courtwatching program implemented, and so forth. Moreover, despite "feeling overwhelmed by immediate tasks" (Levitan, Field
Notes Dec. 1978) staff were actively developing new ideas about things they wanted to initiate.

But, before going off on any new tangents, they wanted to collect a data base, an information resource which would give them a clearer assessment of what had been done before moving in new directions. However, they seemed to have no effective way to measure the work of the first four years. There had been various changes in their original program, but had these changes been "good"? Were the new materials being used? What should be continued or discontinued? The staff considered it crucial to do a current assessment of needs, continuing and new, to find out what needs were actually being fulfilled, and what gaps remained. As well, they considered it important to develop an ongoing means for documentation to facilitate regular assessments of the project.

There were, in addition, other basic reasons for initiating an evaluation of the project at that particular time. There were certain political pressures on the project because of the impending amalgamation of Legal Services and Legal Aid. This had implications for the structural location of the project, since at one time the Attorney General, Garde Gardom, had indication that educational programs should be in the Ministry of Education. It was felt that the position of the Schools Program within the new LSS might be in some difficulty unless linkages with other components of Legal Information Services could be clearly and effectively assessed.

This combination of factors—work overload, lack of operational program objectives, absence of an adequate assessment of existing materials and services, some insecurity about the structural location of the program, and a strong desire to implement some new ideas for the program—formed the basis for the decision to do an evaluation. From the staff's perspective, the evaluation was not intended to tell them whether they were "good" or "bad".
Rather, it was agreed on as an effective tool for use by a working group which had developed the confidence to undertake some form of organized assessment of its own operation. Staff felt that their recognition of the need for an evaluation was a positive move generally. It was a sign that the program had grown and required several new modes of accommodating the existing and new needs of its clients (Levitan, Field Notes December 1978).

Having decided to do an evaluation but unsure of how to proceed, the staff of the Schools Program contacted me to "refine and clarify what the Project staff had in mind" (Levitan, Memo to Schools Project January 1979). Was the evaluation going to be designed to meet the internal needs of the program staff or was it to be set up to measure the effectiveness of the program in terms of some set of external criteria? After a series of intense discussions and negotiations, which took place between the director of the Schools Program, two staff people, a secretary, a half-time person responsible for sending out publications, and myself, the decision to conduct an "internal evaluation" was reached. A staff member later wrote: This was "one of the most important aspects of the evaluation because it must be very clear that everyone understands and has the same expectations to avoid feeling threatened in any way and to ensure full participation and cooperation so that everyone gains from the process:"

As Sheila Fruman, Project Developer of the Schools Program, described this process (from her vantage point at the end of the evaluation study), the decision to conduct an "internal" evaluation had implications for the kind of report produced as well as for the role of such a report.
Often an evaluation report appears some time after the evaluation research has been completed and consists of numerous recommendations for changes based on that research. Only at that point do staff members discuss the recommendations. This report, however, comes not only after completion of the research but also after a thorough analysis of the research findings and the subsequent implementation of changes in both program policy and activities. It doesn't call for change in the traditional sense of such evaluation reports, but rather documents change. ... We chose this method because we wanted to avoid a situation in which program staff are confronted with an "outside" evaluator's report, months after the need for change is recognized, with recommendations for major changes in several areas at the same time. Such a situation can be overwhelming and, indeed, threatening, and often results in many of the recommended changes not being implemented (Preface to the Evaluation Report, 1980).

Instead, staff was to play an active role in the formulation of research questions, data collection and interpretation. The emphasis was "not to be on a written document but rather on the process of gathering data, interpreting it, and arriving at consensus decisions on the changes needed." Staff stressed their need for information that could be assimilated and used fairly rapidly because they wanted to start implementing suggested program changes during the coming fiscal year (1979-80).

Two further consequences resulted from this decision: one was that I remained as a consultant to the evaluation rather than shifting to the position of evaluator. This meant that staff and the consultant met together as a group to decide what specific information would be required, how to divide up the work, what the timing would be, and who would be responsible for analysis and interpretation of the results from different sources. As well, our results were communicated to the LSS senior staff and members of the Board. Each part of the evaluation process was therefore fully documented in memos, transcripts of meetings, research and so forth, so that subsequent production of the evaluation report could be as clear and direct as possible. The second consequence was that production of a final report was assigned subordinate priority to the actual process of doing the evaluation.
The Internal Evaluation: Evaluation Needs and Methodology

The decision to do an internal evaluation of the Schools Program was reached in February 1979. By May, we had evolved the following set of needs that were considered basic to program change. These were:

1. **The need to evaluate various existing program components:**
   This necessitated information on the type/quantity/quality of present curriculum materials, the newsletter, the courtwatching program, and workshops.

2. **The need to evaluate users of School Program materials and services:**
   Staff had to have a much better idea as to who their users were, what were the most important needs of these users, and what might be feasible for the Schools Program to provide (revised and/or newly developed programs) to meet those needs.

3. **The need to find effective means for increased involvement with teachers and community resources:** This required that staff find ways of making more direct contact with teachers, staffs of Community Law Offices and Commission-funded libraries, and other community groups interested in improved legal education in schools.

4. **The need to create more effective program integration:** Staff indicated a strong desire to go through the whole program design to see how the parts did (or didn't) relate to each other, and how policy and procedures could be developed to produce greater coordination of components.

5. **The need to improve policy creation/implementation in curriculum development:** This process entailed the increased knowledge of teacher preferences along with the creation of new funding policies and procedures for the development of curriculum materials (print, AV and curriculum packages).
6. The need to develop an improved distribution system for curriculum materials: Staff expressed a strong need for accurate, up-to-date information on inventory, sales, and product evaluation so as to improve and extend the system for distributing materials to schools throughout the province and elsewhere in Canada.

7. The need to match staff skills with new program priorities: This process meant some revision in the distribution of responsibilities of staff members and changes in job descriptions so that these would more accurately reflect the work that had to be done by a specific job holder.

8. The need to evaluate program relationships with other LSS programs as well as other educational/legal groups and institutions: It was crucial that staff explore their interconnections with other parts of the LSS and the quality of their relationships with other groups and institutions that play significant roles in the general field of legal education.

9. The need to develop ongoing mechanisms for assessment of program effectiveness: Staff considered it necessary to find various ways to monitor program activities so that it would become relatively easy to replicate this evaluation process (systematically and inexpensively) whenever necessary.

There were two basic types of material to be collected as part of the evaluation: 1) various lists of facilities (Secondary Schools, Community Law Offices, Commission-funded libraries) and receivers of services such as the newsletter and workshops which could be mapped; 2) information from documents, annual reports, memos and from interviews in a teacher survey. The decision to survey teachers, to ask them "what they prefer before making any big change" in either the content or delivery of project materials and services (Levitan,
Field Notes May 1979) fit in with the generally expressed need to involve teachers more directly in all phases of the program's work. Staff noted that it would be a way to "make direct contact with teachers, in person, and (hopefully) on an ongoing basis, not just on paper."

The survey was designed to establish: 1) what was the level and type of training of those who teach these courses; 2) what were their course objectives and classroom methods; 3) what were the major sources of information used in their classrooms; 4) what did they know and use among the variety of materials and services offered by the Schools Program; and 5) what new/improved materials and services would they appreciate having available for use in the future.

Since it was too costly to interview teachers in all districts, a sampling technique was decided upon. In addition, funding for the hiring of a summer student was made available through the "work in Government Program", and the program hired Les Berkes, a graduate student in Legal Anthropology, to assist with the research.

The teacher survey went through the following phases: the interview schedule was designed and pretested; the sample of districts and schools was chosen based on consideration of location (rural, urban, suburban) and population of Law 11 and GB 12 classes; and letters were sent out to arrange convenient interview times. Then, between May 23 and June 27 1979, Sheila Fruman, Projects Development Officer of the Schools Program, and Les Berkes interviewed 53 Law 11 and General Business 12 teachers. The interviews were conducted in the Terrace/Smithers, Fort St. John/Dawson Creek, Prince George/Williams Lake, Victoria/Saanich, North Vancouver/Langley, and the Kamloops regions. The areas were selected mainly on the basis of their geographic distance from the Lower Mainland and the presence within each area of Community Law Offices and libraries that had received funds from the LSC. The number of Law classes at each school within each area was also taken into consideration.
In addition, the consultant did interviews with other LSC staff who worked directly with the program, as well with past directors of the program. There was also a suggestion that the evaluation interview a sample of students in Law 11 and General Business 12 classes around the province. Staff knew that this group of users were less directly accessible than the teachers but obviously represented the largest target group benefitting from their products and training efforts. However, in light of the enormous expenditure of time and resources required to do this survey, it was decided to limit this research to a review of enrolment levels (for 1977-78 and 1978-79) as an indirect indicator of student interest in law ans law related subject matter. Finally, an "ecology of materials", was developed through a process mapping the location of classes, workshops, and Community Law Offices, and analyzing the extent of overlap among the facilities, services and materials use throughout the province.

Implications of the Internal Evaluation and Its Methodology

The fact that staff were fully involved in the design and implementation of the evaluation had several implications for them and for those who use their materials and services. Some of these became evident very early in the evaluation process. In particular it was apparent almost immediately that the process of data collection served two functions: it provided a data base; but, in addition, it provided an active means of establishing personal contact between teachers and program staff. Through the process, program staff gained a greater understanding of the resources and limitations for teaching law in several communities around the province. For example, does having a Community Law Office in the area make any difference? What about a Community College? What about a high proportion of native people living in the area? Staff gained a more accurate picture of the needs of law teachers working in
different types of community environments. As one staff member wrote: "All of us have a better idea of who our users are, what they need, and what it's possible for us to provide."

Staff were also able to gain an understanding of situations as the evaluation progressed. Information was exchanged and analyzed as soon as it was gathered. This rapid and open form of communication had major implications for the success of the evaluation. Staff were able to develop an analysis by jointly working through the information on a particular component of the program. It was possible to implement changes quickly rather than waiting until the end of the "evaluation period". There was no problem of anticipating, with apprehension, the appearance of a formal report with numerous recommendations constructed by an "outside" or non-program person.

By doing much of the day-to-day work of the evaluation themselves, all members of the staff developed the skill and ability to understand the scope and limitations of that data. Also, by retaining control over the collection and interpretation of the data, staff were able to decide what they really needed, in what order of priority, and how it could and should be used. They did not have to rely on the interpretations of an "outside evaluator"; but the presence of an independent consultant provided an "outside voice". The consultant worked to retain a certain amount of distance from the evaluation process and was willing to ask the hard, critical questions which staff had to be prepared to answer to their own satisfaction. After all, it was the staff of the Schools Program which was going to have to live with the consequences of the changes that they implemented.
Subsequent Aspects of Program Change

1. Creation/Implementation of New Policies

In connection with the various components of Schools Program operation, it is important to recall that staff emphasized the need for greater overall program integration and improved policy/procedures in the area of Curriculum Development. As each component of the Program was researched and criticized Project staff and the consultant explored various new strategies for the coordination of activities and integration of efforts around a series of agreed upon priorities.

This was spelled out most fully in the area of curriculum development. For example, project emphasis was given to the promotion of existing or recently developed materials rather than starting to fund an entirely new set of productions immediately. This was done in an effort to focus energy on existing needs and not allow staff to spread themselves across too many new and demanding tasks.

Also in curriculum development new policies for funding were created which were tied in with yearly topic priorities. There were two different types of curriculum materials funded: 1) major projects and 2) locally developed projects. When the Schools Program began there was an emphasis on locally based materials. But since the survey indicated that teachers looked to the Schools Project as a major producer, staff felt they wanted to have some more direct influence on content and quality of materials. Staff members moved toward a policy of greater control over what is produced and made available provincially. This shift still made a specific part of the budget available for high quality locally developed materials.

The move toward greater integration was also reflected in a new emphasis on the development of comprehensive curriculum packages. These packages contain the background, history, legislation, case materials and
other information related to a specific topic, along with an up-to-date bibliography. Such a package could be used as a resource for a workshop and publicized in a newsletter published just before the workshop. This was another way in which staff members began to create more effective coordination of their efforts.

2. Restructuring Program Responsibilities

As the research phase of the evaluation process drew to a close, staff and consultant were able to focus their energies on producing a more accurate outline of the emerging set of program components and related job descriptions. At the end of October 1979 Schools Project staff constructed a revised Program Design which included a more comprehensive statement of purpose and a division of program activities into five major areas: Curriculum Materials; Education Programs; Newsletter; Liaison and Evaluation. Then the actual job descriptions of staff members were changed to match individual skills more closely with the new program design.

3. Developing Ongoing Evaluation Mechanisms

In an effort to ensure that the Schools Program "is responsive to and reflects the needs of teachers and students in B.C.", staff set up a series of procedures for program monitoring. Both the Teacher Survey and the Newsletter survey were designed to become annual events. In addition, staff put in place new methods for recording and collecting information about various other aspects of the Program. These included: 1) the new inventory system for all print and A/V materials; 2) evaluation forms for workshop participants and organizers; 3) evaluation forms for all A/V materials (teacher feedback); 4) forms to document yearly enrolment in Law 11 and General Business 12, and 5) maps to show locations of secondary
schools, workshops, Community Law Offices, LSS funded libraries, etc. All these were developed to provide the raw material for continued scrutiny of Schools Program users, materials and services on a monthly/yearly basis. Analysis of this information should enable staff to regularly assess program directions and make changes when necessary.

Assembling the Evaluation Report

Since this was a collaborative process, my drafts, which began with a history of the program turned in on 7 September 1979, were all completed and given to Pat Robertson, the staff person in charge of print media, on 13 February 1980. These drafts were altered and changed during the succeeding months. In May 1980, a second draft of the entire evaluation report, edited by Sheila Fruman, the Projects Developer, was circulated to the Directors of the Schools Program, Legal Research, Public Legal Education and Library Services and to me for comments. In her memorandum of 15 May 1980, Sheila remarks that it is an 'amalgamation' of sections written by all the staff members and by me.

In mid-August 1980, I reviewed the third and final draft of the evaluation report just before it went to the printer. At that time I noticed that the "Results" section which in my initial draft had been 11 pages, had, in that final typewritten draft, been reorganized and substantially expanded to 21 pages. It presented additional program changes, and documented in greater detail just how the changes had been accomplished in each component of the program. It was also better written than my initial draft, of course. This indicated the degree to which the Schools Program staff had become "experts" regarding their own evaluation and organizational changes.

In her preface to the report, Sheila Fruman raised the question, "Why do a report at all?"
When we realized that many of our original questions had been answered through the process of evaluation and that the required changes had been implemented, we were not sure there was a need for a report at all. Our internal needs had been met, after all, and since it was an evaluation for ourselves, what additional need could there be for a report?

However, the Schools Program staff realized that there was "value in documenting both the process of the evaluation and the findings for the following reasons":

First, as the evaluation progressed and we became aware of the original history of the Program, we gained a different perspective on what we were doing. We also decided that a report documenting such information would be useful for future reference, both for Program staff and others involved with the Program, such as staff of other departments of the Legal Services Society and Legal Services Society Board members.

Secondly, the field of legal education in Canada is a relatively new one. Many provinces have no programs whatsoever and those that do are small and isolated. We felt a report could contribute to the broader development and growth of legal education through sharing our experience with others.

Finally, at the present time in Canada, the evaluation of legal education programs is virtually non-existent. This report is our contribution to this uncharted territory.

The intent of this report is not to tell you whether the Schools Program is good or bad, or whether the Program should continue or be discontinued. It does not pose basic questions about the value of legal education or the need for a Schools Program. We feel, however, that many of these questions are answered implicitly in this report (Preface to the Evaluation Report 1980).

In subsequent months the Schools Program Evaluation Report 1980, was requested by various groups involved in legal education programs in other parts of Canada, the United States, and other Commonwealth countries.
5.4 The LSS Schools Program Evaluation: The Locus of Evaluation Ownership and the Process of Organizational Change

Basic Areas of Negotiation

The evaluation of the Schools Program involved a small number of direct negotiators—the program director, the two staff members, the secretary, the summer student and the consultant. None had any experience doing program evaluations, except me, and I was hired specifically for my skills in that type of consulting. All the permanent members of the program were experts in their own particular jobs. They were able to provide the orientation to program activities and substantive details required to take advantage of my knowledge and skills, as well as those of the summer student trained in legal anthropology.

There were multiple linked negotiations in the process of this evaluation. As indicated previously, the evaluation process developed from a series of repeated discussions in which the work was clarified, alternatives were explored, and my skills were put to many different uses. It was an extremely democratic working arrangement and there was a fairly equalitarian division of labour. The pattern of interaction tended to involve an ongoing exchange of insights, information, and analytical suggestions.

The negotiators had complementary stakes in the evaluation process. Essentially the staff of the Schools Program wanted to emerge from the evaluation process with a blueprint for future program planning as well as a significant reorganization of their existing commitments and activities. My stake in the process included increased understanding of the operation of an education program, insights into aspects of public interest law, as well as the fee (altogether just over $6,000) I earned for my consulting work. It offered me the opportunity to work closely with a program that I respected and valued because
of the quality of services and resources it provided in the school system in British Columbia.

Of the four case studies examined, the transactions in this situation were the most visible. There was little need to hide difficulties or lack of knowledge regarding various aspects of the research process. It might be said that we all had a vested interest in being as "upfront" as possible so as to maximize the exchange of information and skills. Virtually all organizational issues were made visible, and most were negotiable. For example, I was encouraged to become directly involved in the process of financial planning and budgeting for the program. The provincial government had brought in the policy of "zero-based budgeting", but was not supplying the various Ministries and programs with adequate information about the procedures for filling out these new budget forms. When this problem arose at the Schools Program, there was never any attempt by the director and staff to hide their confusion about how to deal with these new government requirements. I suggested that we ought to be able to use the data and analysis procedures already available from the evaluation process to assist in the construction of the new budgets. We agreed to work collectively on the problem and it was then possible for us to provide the priorities and financial rationale necessary to construct the numbers and fill in the boxes on the budget sheets. This exhibition of fiscal competence permitted the director and staff to maintain a greater measure of control over the budget allocation for their program for the coming year.

There was one significant issue over which there was little or no overt negotiation during the time I worked as consultant to the program. This issue related to the dissatisfaction in the relationship between the two main staff members and the director. In separate interviews with all three they expressed various concerns over the unclear division of labour, ideological conflicts, levels of competence, and working styles. However, neither the director nor
the staff members were willing to confront these issues, or the underlying problem of the balance of power and responsibility within their small group. Since I had been hired as a resource person for evaluation of the program (as the unit of analysis), I made little effort to encourage such a confrontation. However, I did suggest that they might want to hire someone skilled in assertiveness training and conflict resolution to assist with their ideological and interpersonal difficulties. This was never done. In some sense the tasks of the evaluation process tended to redistribute work responsibilities and this permitted the director and staff members to defer/deflect a possible confrontation. The balance of power and responsibility within the negotiation context of the program was very significant, but it was always kept beyond the boundaries of the negotiation context of the evaluation.

Both the staff and program director were involved in all the basic areas of negotiation including: the decision to employ a portion of their resources on an evaluation; hiring the consultant; exhibiting an early and continued interest in the design and implementation of the research plan; involvement in interpreting the results and their inclusion in setting new policies and making rapid organizational changes; and finally, control over the assembling, production and distribution of the evaluation report. (The staff of the Schools Program exhibited a higher level of involvement in the evaluation process than did the director.) Ownership of the process and the product was therefore located firmly within the Schools Program. This pattern of ownership may be interpreted as the source, cause or context for the relatively quick and successful series of organizational changes implemented within the Schools Program during the two-year period of the evaluation study. It does suggest that where there is a relatively high level of involvement of negotiators across all the basic areas of negotiation, there is greater potential for the
evaluation results to be rapidly integrated as part of a process of organizational change.⁹

In addition, their involvement and control over the product of the evaluation, the 1980 Evaluation Report, increased the potential readership of that document. The program staff used preliminary findings at a Conference on Youth and the Law (Saskatoon, May 1980) and requests for copies were received from Legal Education Projects across Canada and in the United States and Britain. The evaluation report was also used as a resource to persuade/educate groups and organizations within British Columbia regarding the scope and possibilities of legal education in the school system, e.g., with members of the LSS Board of Directors and the executive of the British Columbia Bar Foundation.

The Impact of Structural Context on Long-Term Change

The comprehensive evaluation process and the subsequent organizational changes in the Schools Program were accomplished with a minimum of disagreements and a strong sense of the mutually productive collaboration taking place. However, many indirect effects resulted from factors which originated in the structural context of the Schools Program and tended to undermine the achievement of substantial long-term changes in program operation. Essentially, these obstacles to permanent changes in the management and scope of the Schools Program lay in the strongly hierarchical and instrumental approach which characterized the upper level of management of the LSS and its Board of Directors.¹⁰ Their managerial priorities continued to be the control of subordinate levels in the organization and a shift of resources to Legal Aid and away from Legal Information Services.

As indicated in the organizational charts in Section 5.3, the Schools Program was one of several units which make up Legal Information Services. There had never been a particularly well-developed coordination among these various LIS
units and management of the overall LIS Program had been weak. There was no systematic process for long-term planning, implementation and evaluation for the LIS Program. Moreover, despite pressures and lobbying from some units, there had been a policy of "no response" at upper management levels. All efforts to introduce democratic/equalitarian forms of decision-making had been disregarded. Much of the dissatisfaction at lower levels of the organization resulted in a strike during the spring of 1981. These problems within the structural context of the Schools Program had varying effects within the negotiation context of the program itself. For example, despite their strong commitment to the evaluation process and a clear sense of ownership of the evaluation product, the staff and director never made efforts to resolve their central problem regarding program management, i.e., was it to continue as a form of ambiguous hierarchy or would it shift to being a democratic/equalitarian work group.

This problem was resolved in August 1981, by a decision at the LIS level. That is, the director of LIS replaced the previous director of the Schools Program (who resigned to go back for an advanced degree) with a person who shared the managerial ideology and practices of the upper level of management and the majority of the Board of Directors of LSS. Whereas the previous director of the Schools Program was a woman with a teaching background and liberal politics, the new director is a man with a teaching and legal background who clearly exhibits a strong individualistic, conservative approach to politics and education.

Recently a staff member who is on leave remarked that the Schools Program would undoubtedly continue, but it seemed difficult to go back since LIS had repeatedly failed to meet the challenge of providing some direction to the program. However, from the standpoint of LIS and the upper levels of LSS management, perhaps they are finally satisfied that they now have a director of the Schools Program who will control the program and produce safe, reasonable materials and services.
Overview and Implications for Knowledge Production and Use

This final case study provides the description and analysis of a relatively unusual process of internal program evaluation, where research findings were immediately fed back to project staff for discussion and interpretation, and organizational changes preceded the production of any formal evaluation report. The evaluator remained an external consultant throughout the process, suggesting a composite research strategy which contained consensual elements (using existing data on enrolments, materials in print, and descriptions of teacher workshops), contentious elements (questioning the ongoing funding of the Courtwatching program, constant funding of new curriculum materials, and the absence of any overall plan for sales and distribution of program materials), and a limited amount of paradigm-challenge (trying to get staff and director to push for a more comprehensive process of policy planning throughout the LIS component of the Society). In this program context members of the staff and the director were fully involved from the initial decision to allocate resources to an evaluation to the creation and implementation of new policies and the redistribution of program responsibilities. Staff were more involved than the director, particularly in creating the research design, data collection, and assembly and production of the report. However, the director was active in using the evaluation document in promoting the achievements of the program with the Ministry of Education and the Board of Directors of the LSS.

There never was any direct involvement of the upper levels of LSS management or the board in any aspect of the evaluation process. This simplified the negotiations entailed in accomplishing the evaluation research and subsequent changes in policy and program components. But it meant that these changes were made in isolation from the realities of the structural context of the program. Moreover, during the evaluation the central issue of the distribution of power
and authority within the program was defined as being outside the legitimate boundaries of the evaluation process.

Therefore the evaluation process and the program changes it set in motion could be easily stopped, reversed or redirected by the newly hired director whose management perspective and politics conform much more closely to that of the upper management and the majority of board members. Since the moves toward democratization and collective management which characterized the evaluation phase of the program were implemented in only one other organizational unit, the Legal Resource Library, not even the complete LIS Program could be persuaded/educated/negotiated into a shift in management policy. Consequently, there was no possibility that changes in the negotiation context of the Schools Program could affect any shift in the structural context of the LSS.

Therefore we can recognize somewhat similar characteristics of structural context as offering obstacles to long-term change in this setting as we observed in the JFSA evaluation. In the case of the Schools Program there was an ongoing but unaddressed conflict between the director and staff regarding the balance of power and responsibility within their small group. This conflict did not interfere with the negotiations involved in the evaluation process (as did the overt conflict between the board and director in the JFSA assessment), however it tended to prevent development of a carefully planned campaign for the hiring of a new, ideologically compatible Schools Program director when the original director decided to resign.

In addition, there was a parallel problem here regarding the appropriate unit of analysis for the accomplishment of lasting organizational change. Here too, the appropriate unit for evaluation and subsequent change should have been larger; most natural in this context might have been the entire LIS component of the Society. The research strategy in an LIS evaluation would have encouraged participation from the "enemy", the staff of Legal Aid and members of the LSS
board, and offered some opportunities for negotiation regarding general policies on organizational management, control, and resource allocation in the larger context. This case study provides another example of the need to determine an appropriate unit of analysis and the boundaries or limits to organizational change as part of any model of knowledge production and use in the policy process.
1. There are several other options. However, the choice to do an evaluation may be related to some combination of organizational strength, insight into the problem, and desperation.

2. She was a single parent with five children, a very good professional background, and a network of supportive friends in the Jewish community.

3. This involved editorials in the Jewish Western Bulletin, letters to the editor, going to just about every Hadassah meeting in town, lobbying for a higher profile for the agency and this did pay off in a substantial increase from Jewish Community Fund and Council.

4. They have implemented certain kinds of recommendations, including the development of a loan fund, called the Hebrew Assistance Association which was established in 1980, and in this past year the JFSA has begun to do outreach work at the Talmud Torah, the Hebrew day school, where there are increasing numbers of kids from single parent households, kids with emotional problems, and kids with learning disabilities.

5. The program was called the Public Schools Legal Education Project when it began. In fall 1977, the name was changed to the Schools Legal Education Project in line with the Ministry of Education's revised policy of funding private schools. In the fall of 1979 the program changed its name to its present title, the Schools Program, Legal Services Society. The name Schools Program was used throughout the report for consistency to refer to all phases of the program.

6. There were several sources of data that were available to me in order to put together this case study. There were, of course, taped and transcribed interviews with all members of the Schools Program staff, the present and all preceding directors, as well as the staff and Director of Legal Information Services. The evaluation process was well documented and could be reconstructed according to my own worldview and framework for analysis using a series of memorandums, reports, survey results, and my own field notes. In addition, from 1 May 1979, through 10 August 1979, all our major work sessions had been taped and were transcribed by the Schools Program secretary. There was one other source of data I should mention here, i.e., my blue account book. Since I was hired on an hourly basis, and billed the Program each month for my services, I kept very close track of what I was doing and how long I worked on any task. The account book had headings for the date, the type of work done (and with whom), the time, and the fee. Therefore, it was relatively easy to produce my billing statements for the Program, but it also had the unanticipated consequence of providing me with a running record of the stages through which the evaluation and implementation process proceeded.

7. A model for this kind of collaborative method used on a large scale is provided in Rowbottom (1977).
8. Zero-based budgeting requires an analysis and rationale for each budget item. Nothing can be carried out automatically from preceding years.

9. A similar pattern of interactive methodology, work setting ownership of the evaluation, and organizational change is reflected in the study of Vancouver Women's Health Collective done by Kleiber and Light (1977).

10. This traditional approach to organizational management is presented in Drucker (1974) and Robbins (1980). For alternative perspectives in work organizations see Greenberg (1975), Bernstein and Bowers (1977), and Crowfoot (1981).

11. During May 1979, I did some consulting work for Penny Bain, the Director of LIS. This involved a brief analysis of the Public Legal Educational Program in order to develop a new definition of the job of director of that particular program, as well as an outline of a data format required for adequate program reporting by projects funded under the PLE budget. Again in September 1979, through January 1980, I did occasional consulting with LIS to examine the feasibility of doing an overall LIS evaluation and to develop a new and more coordinated set of funding guidelines for PLE that would ensure greater consistency in project development and support. The overall evaluation of all the programs in LIS was never done, but in 1981 the staff of the Library Services did their own evaluation, with the assistance of members of the Schools Program.

12. During the months of October-December 1981, I did some additional consulting for the Schools Program. In this situation I assisted members of the Program staff and the Director to prepare a brief for the Bar Foundation of British Columbia which outlined a proposal for an expansion of legal education in the school system throughout the Province. The Bar Foundation had put out a request to all the organizations concerned with the legal education in an attempt to develop a funding policy for the distribution of the increased revenue from trust accounts which it supervises.
In recent decades the growth and co-optation of social reform programs into the structure of the State, and the parallel development of public policymaking, have precipitated closer linkages between social research and the policy intervention process. There have been two major perspectives employed in the effort to connect social science theory and empirical studies to the formulation of public policy: "policy science" and "policy analysis". Although these perspectives differ in their underlying epistemologies and theoretical statements about the potential uses of policy research, both recognize the publicity and controversy which frequently surrounds policy research, while sharing an optimism about its ultimate usefulness. In addition, both approaches define the evaluation of government-funded programs concerned with social reform as a key area of policy research.

This particular form of policy research, program evaluation, may refer to a wide variety of descriptive and analytical studies of program process and/or program impact. However, there are basically only two models of program evaluation research which tend to be relied upon in the design and implementation of evaluation studies: a conventional model which derives from a positivist paradigm of social research and an alternative model which has evolved from an interpretive (anthropological) paradigm. My research purpose has been to develop a wider selection of evaluation research models which specifically take into account construction of the research "product" and characteristics of the larger structural context in which such products are designed to be used. The set of models outlined in Section 6.3 of this chapter are the final result of a theoretical and empirical research strategy
which focused on the relationship between the structural contexts of program planning and implementation and the negotiation contexts of the production of program evaluation documents.

6.1 The Theoretical Approach and Relevant Debate

My theoretical strategy has involved an effort to relate the more stable structural characteristics of the social, political, economic and organizational context of reform programs to a series of six theoretical/practical issues in the evaluation process. These issues included: delineation of all the major actors; organizational placement of the program evaluation work; choice of the general research strategy; selection of an appropriate research model and methodology; the construction and content of the research report; and the planning for research utilization.

The Analytical Strategy

The first step in this strategy involved a critical review of the two major models of program evaluation research described in the contemporary literature on policy research in education. I examined each model in terms of its main assumptions and directives regarding resolution of the six theoretical/practical issues. Comparative analysis of the two models indicated that they differ on virtually all six issues in policy research. The positivist evaluator tends to focus on abstract features of program output or impact which are relatively simple to measure; while the interpretive evaluator emphasizes the processual components of program operation and service delivery as they can be ethnographically discerned through observations and interviews. Furthermore, exponents of each approach are cognizant of the strengths of their own model, but often find it difficult to recognize:
a) the limitations of their preferred research model, and b) the complementarity of the two models for comprehensive evaluation studies. Finally, evaluators seldom mention the extent to which both models minimize the significance of the larger political/economic context of program development and implementation in shaping the evaluation process. Thus, an expanded repertoire of models for knowledge production and use does not necessarily require a search for new designs and methods, but more extensive scrutiny of the empirical settings in which evaluations are accomplished and more comprehensive analysis of the structural contexts in which they are requested, funded, supervised and used.

The second phase of the analytical strategy entailed construction of a revised framework for the analysis of program evaluation research which would extend beyond the limits of existing models. The overall framework presented relies on the theoretical approach of symbolic interaction, and more specifically on the negotiated order paradigm developed by Anselm Strauss (1978). He describes the social order in terms of the process of ongoing negotiations carried on by many types of participants, each with distinctive world views. Negotiations are conducted simultaneously at several levels in the social order, and each level can offer resources to and impose constraints on others. The two key terms in his analysis, "structural context" (referring to the more stable, salient structural properties that bear on negotiation) and "negotiation context" (referring to the differing contexts of interaction among negotiating parties and the properties of those contexts), provide the general categories of description and analysis in my theoretical framework. Using several aspects of Strauss' theoretical argument, as well as other sources (Rein 1976; Wiseman 1979), I construct a framework for the analysis of program evaluation studies which:
a) specifies the social, political, economic and organizational components of the structural context of program planning and implementation which bear on the policy research process generally, and program evaluation in particular, and b) presents the six theoretical/practical issues of policy research as basic areas of negotiation in the program evaluation process.

This framework provided the theoretical perspective for the description and analysis of two sets of case studies discussed in Chapters 4 and 5. However, before turning to a summary of those case studies and the results of my analyses, there are certain questions regarding my choice of theoretical framework and a portion of the debate regarding the negotiated order paradigm which must be addressed.

The Choice of Analytical Framework and Relevant Debate

I want to outline briefly why I considered the negotiated order paradigm particularly suitable to an analysis of program evaluation. One significant strength of Strauss' negotiated order paradigm is that it places the actual negotiations at the center of the analytical model but does not exclude our consideration of alternatives to negotiating—persuading, educating, manipulating, appealing to the rules or to authority, and coercion (1978:x).

Through use of the paradigm we are directed to examine the linkages between the structural context in which programs are developed and the negotiation contexts in which they come under the scrutiny of program evaluators and policymakers. Another basic strength of this approach is that the negotiated order paradigm permits us to assemble an array of research data which provides: a) important insights into patterns of social order and social change in programs developed under the existing socio-economic framework in North America, and b) the initial conceptual tools for developing alternative
practices for the planning and implementation of welfare/housing/educational/health programming in other types of socio-economic settings.

Debate and criticism of negotiated order theory (or 'perspective' or 'approach') has centered in the work of Benson and Day (1976) and Day and Day (1977). Maines (1977) deals with some of the criticisms, while Strauss (1978:247ff) handles some of the typical misapprehensions or misunderstandings of the position, as well as presenting some answers to those diametrically opposed to the negotiated order position. I want to review very briefly here, selected critical objections and answers to these objections which have emerged from the work of this dissertation.

**OBJECTION:** Negotiated order research seems to apply essentially to microscopic interaction rather than to macroscopic-structural problems.

**RESPONSE:** All of the case studies involved analysis of negotiations carried on among units of varying organizational scale, up to and including departments of Provincial governments and Crown Corporations; in terms of temporal scale certain evaluation processes extended over months and years.

**OBJECTION:** Negotiation order researchers assume that social order, or organizational relationships, are unstable, ever changing.

**RESPONSE:** There is no need to debate the differential stability and instability of specific social orders, e.g. some programs were designed to have a certain life span (SRP and NIP/RRAP) while others were structured as continuing programs (JFSA and the LSS Schools Project) where negotiation contributes to maintaining program stability and minimally disruptive change. In the program evaluation process it is crucial that negotiators try to anticipate those recommendations for organizational change which will tend to maintain the stability of positive aspects of program structure and delivery, while suggesting program transformations which will improve aspects of day-to-day operation and long term impact.
OBJECTION: Negotiated order theorists tend to emphasize cooperative relations and neglect relationships which focus on coercion and conflict.

RESPONSE: All four of the case studies of program evaluations presented here have been shown to include varying patterns of cooperation, conflict and tactics which implied coercion. In some instances there was overt antagonism exhibited at every stage in the negotiation process and among all the parties to the negotiations (e.g. NIP/RRAP evaluation), while in another case (LSS Schools Project Evaluation) there was a very significant pattern of collaborative negotiation among the immediate parties to the evaluation process, but the overall evaluation was accomplished in isolation from the powerful elite of the larger structural context whose political/managerial priorities were totally antagonistic to those of program members. In the latter case it became evident that coercion was essential to relations among negotiating parties in the larger organization.

OBJECTION: Negotiated order may be taken as a complete explanation of social order.

RESPONSE: Following Strauss' position (1978:250), I have made no attempt to show that negotiations could explain every aspect of the processes and products of program evaluation. What I have emphasized is that no program evaluation, no matter how systematically/scientifically/rigorously designed has ever been implemented in a field situation without a substantial set of negotiations in several basic areas (e.g. hiring, conceptual/methodological frameworks, data analysis, document assembly) in order to accomplish the evaluation document which decision makers have contracted for and evaluators have constructed to meet those initial specifications.

OBJECTION: The negotiations perspective deals with the "subjective", naively accepting the perspective of participants under study.
RESPONSE: This objection was handled in Chapter 2 where I addressed the criticisms levied against all research models which include interpretive materials as essential to the process of description and analysis. However, I do want to stress here that use of the negotiated order paradigm developed by Strauss requires that we take into account both actors' viewpoints and structural considerations. I would agree fully with Strauss when he asserts that "any and all sources of useful data should be tapped" (1978:251).

OBJECTION: Not everything is negotiable, as negotiated order analysts claim.

RESPONSE: Again referring to the four case studies, my analysis showed clearly that in various settings where program evaluations were accomplished at some structural distance from the center of policy planning, no amount of local negotiating could break down the resistance to change and the relatively fixed policy intervention process which characterized the larger social structure. Strauss points out that "not everything is either equally negotiable or--at any given time or period of time--negotiable at all. One of the researcher's main tasks, as it is that of the negotiating parties themselves, is to discover just what is negotiable at any given time" (1978:252).

The problem of determining just what is negotiable at any given time suggests that we must constantly address the topic of "limits" and their discovery. Use of the Strauss paradigm has permitted me to describe and analyze both the opportunities and constraints presented by the social, political, economic and organizational aspects of structural context. For example, the definition of a "social problem", the legislative mandate of programs designed to deal with the problem, the budgeting and organization of service delivery in those programs, and the managerial ideology of those who run the programs--all these may enhance or impede the process of program evaluation.
Within the context of each of the program evaluation case studies, I have been able to describe and analyze limits regarding: the number and representativeness of major actors, possibilities regarding organizational placement of the program evaluation work, options and consequences of choosing a particular research strategy, the appropriateness and acceptability of the research model and methods, the construction and content of the research report, and the extent to which research utilization could be planned and facilitated in advance of the completed report.

Nevertheless, sociologists who follow the theoretical approach which emphasizes political economy and political power (mainly the neo-Marxists) continue to have major disagreements with the negotiated order paradigm. Generally it is argued that negotiation is contingent on organizational structure, i.e., that it is the politically and economically ordered characteristics of groups which yield the structural conditions under which negotiation does/does not occur. Thus in discussing Canadian public policy Mahon (1977) follows a deductive approach which begins at the macrolevel of the State, emphasizing the way in which the State is structured to perform its role as the organizer of hegemony and the system of representation inside the administrative apparatus. This in turn affects the ideological position of the dominant segments of the civil service and shapes the patterns of knowledge production and use in policy settings (1977:175).

Part of the work of this dissertation has been to explore the nature of the relationship between structural context and the negotiation contexts of program evaluations without taking the completely determinist position of Mahon and others (Panitch 1977). It would be reasonable to use the analytical framework developed here and "expanded" to the level of the Federal administrative structure to describe and analyze the nature and extent of
control exercised by key departments such as Treasury Board and Finance. In the concluding section of this chapter I have outlined the proposed new controls over program evaluation studies presented by Treasury Board. The research question then involves the extent to which these directives will be followed by the civil service, the variety of interpretations they will be given, and the models of negotiated knowledge production and use which will be implicated in subsequent shifts in Canadian public policy.

6.2 Overview Of The Case Studies: Implications for Knowledge Production and Use

Studies in Housing Policy

In Chapter 4 I described and analyzed the structural and negotiation contexts of two evaluation studies commissioned by government to assess delivery of service and program impact during a shift in Canadian housing policy from the initial goal of slum clearance to a new orientation which focused on housing and neighbourhood rehabilitation. The first evaluation concerned the Strathcona Rehabilitation Project, a housing and neighbourhood rehabilitation program located in the Strathcona area of Vancouver. That project was designed to serve as a model for subsequent national legislation regarding the conservation of housing stock and upgrading of community facilities in cities and towns across Canada (Neighbourhood Improvement Program and Residential Rehabilitation Assistance Program, NIP/RRAP). The second evaluation case study was an assessment of the NIP/RRAP program delivery system in British Columbia. It was part of a larger national evaluation of the NIP/RRAP programs originally designed to provide policy research data for changes/amendments to the federal legislation concerned with housing programs.
The SRP Stage II Evaluation: Conclusions

Analysis of the SRP Stage II Evaluation suggested that several aspects of the structural context of program development and implementation had significant impact on the program evaluation process. These included: the consensus model for decision making used by the management committee (the SRC), the large number of actors represented both formally and informally in program management and day-to-day operations, the pattern of resource control and allocation established before the program management structure was fully developed, the constant shifting from formal to informal styles of communicating and decision making, and the protracted nature of all negotiations having to do with the program.

Certain of these characteristics of the program had importance in enhancing or limiting specific areas of negotiation in the program evaluation process. For example, the consensus model of decision making tended to mask conflicts on the SRC (especially between SPOTA and the City), but ultimately some of that same conflict surfaced in relations between the last City representative (indirectly supported by his Engineering Department) and the evaluation research team. The large number of actors formally and informally represented on the SRC contributed to the complex and protracted hiring process and subsequent delays at each stage in the evaluation process. Some of the delays were the direct responsibility of our nine-person study team which wanted to produce a comprehensive evaluation of the program; however, others were the result of having to deliver all correspondence to the four main parties on the SRC and wait for times when all parties found it convenient to meet and make decisions. The formality/informality dualism tended to be confusing to members of the evaluation team and for a long time it prevented us from understanding how to present our case at specific times in the negotiations.
In addition, although the conceptual and methodological strategies employed by the study team were never directly in question, our interpretations of the data were subject to a substantial amount of direct scrutiny. There were negative comments made regarding our "objectivity", and the results of our financial analysis of the actual program budget ran counter to the preconceptions and intuitions of various government representatives and bureaucrats. This resulted in confrontations regarding control over the assembling, production and distribution of the program evaluation document.

The NIP/RRAP Evaluation: Conclusions

In the NIP/RRAP evaluation the negotiation profile was remarkably different, despite similarities in the substantive aspects of the program, some duplication in personnel on the study team, and a complete time overlap in the research and write-up of the report (NIP/RRAP was actually turned in to CMHC Ottawa before the final draft of the SRP Stage II report was handed over to the SRC). The overall evaluation strategy for NIP had been developed by CMHC Ottawa staff and the national coordinator from Toronto before any consultants in the regions were brought into the process. Funding for the evaluation in B.C. was inexplicably split between two research entities (Community Planning Association of Canada--B.C. Branch and Ptarmigan Planning Associates) during a confusing, long distance hiring process that included no identifiable proposal call, interviews or other traditional elements. Authority, financial resources, and the presumed control of research results lay totally in the hands of the NRID research director and the national coordinator, while the study teams in each region were completely accountable for bringing in research data required by these two major negotiators. This assumed balance of control and accountability was directly challenged by
our British Columbia study team in an ongoing series of direct confrontations, as well as in written form in the Methodological Appendix to the British Columbia Evaluation Report.

Based in two substantially different models of research, the conflict between the B.C. team and the National Coordinator was never resolved and the NRID Research Director was never able to mediate or to fire either of the parties. We totally redesigned the interview schedule according to a series of principles which were in direct contradiction to the "testing" approach taken by the National Coordinator. And, due to the richness of detail and ultimate identifiability of the sources of various remarks we did, in fact, retain all original copies of our interviews in order to protect respondents' anonymity. This series of confrontations continued to the close of the evaluation process and was the defining characteristic of the negotiation profile in this case study.

Thus, in comparing the two sets of negotiations, it became evident that in the SRP evaluation despite some ambiguity and risk, along with the major confrontation over the shape and the content of the final report, most of the significant negotiators retained a sense of the novel and challenging aspects of the program which tended to carry over into negotiations regarding the evaluation. The consensus model of decision-making employed in the SRC, flawed as it may have been, encouraged a level of directness and visibility in the evaluation negotiations that was totally missing from the NIP/RRAP case study. In that second instance the process of negotiation in all the basic areas of the evaluation process was characterized by much more covert transactions, where legitimacy boundaries were never clarified and the ostensible balance of power unworkable. This led, ultimately, to negotiations in which conflict and confrontation were the dominant modes of
interaction among the major negotiators. These modes of interaction tended to extend into the negotiations within the work base itself resulting in a lack of trust and agreement ("web of commitment" in Strauss' terms) which could only be overcome by the professional orientation and ethical/political principles of the members.

**Implications for Knowledge Production and Use**

Finally, referring back to my main research concern with the development of additional models of knowledge production and use, there are several conclusions and implications which emerged in these two case studies. First we should note that in both these cases the evaluations were located at some distance from the policy intervention process. There were no information specialists available or involved in the structural context of the programs who were competent to make the necessary connections between the input presented in our evaluation reports and those policymakers concerned with the goals and objectives of housing policy in Canada. Second, both case studies also exhibit what may occur in the direct/indirect confrontation between two models of evaluative research. In the SRP evaluation we saw the implicit outlines of the positivist model in the criticisms presented by municipal level bureaucrats (City Planner and Engineer). In the NIP/RRAP evaluation the ongoing antagonism between the highly structured, controlling and testing approach of the National Coordinator and the more open, searching and interpretive approach taken by the B.C. study team reflect strongly contrasting research paradigms along with the political ideologies implicit in those paradigms.

In addition, there are several implications to be drawn from the individual case studies. Analysis of the SRP Evaluation process suggested that
in research settings characterized by a large number of actors, policy researchers include a preliminary phase in which they quickly study and identify the political, economic and ideological stakes of the key actors before formulating a general research strategy and the model and methods for the study design. The same level of critical analysis may be applied in formulating the research strategy they plan to use; because explicit or implicit, any research strategy has certain identifiable assets and liabilities in specific policy research settings. Finally, easy social relationships between evaluators and key actors in the research setting may provide no guarantee of the acceptance and ultimate use of the research findings. As indicated in Caplan's (1977) study, the utilization of policy research is a complex process requiring powerful intermediaries and a clear conceptual "fit" between the presentation of research findings and the policymaker's orientation to decision-making.

The pattern of negotiations in the NIP/RRAP evaluation process brought out other implications for knowledge production and use in policy formulation settings. The central question which emerged in this context was the location of design and administrative control in the evaluation of nationally funded programs. The series of problems and conflicts encountered by the British Columbia study team in this case study suggested that regional variation in the structural context of programs may be more suitably dealt with through a decentralized process for the development of research strategies, models and methods. It appeared that the control, objectivity and generalizability which are supposed to be the products of uniform design and administration of evaluation instruments may perhaps be "false commodities". If such a uniform product is impossible to generate, then policy researchers
may have the collective responsibility of educating their "consumers" regarding the scope and limitations of the policy research product.

**Studies in Private Social Services and Legal Education**

One case study described and analyzed in Chapter 5 focused on an assessment of the Vancouver Jewish Family Service Agency, a private social service agency which is supported by funds from United Way and the Jewish Community Fund and Council. The second case study concerned the evaluation of the Legal Services Society Schools Program, an ongoing program for the delivery of legal education services and materials in the B.C. school system, supported by funding from the Provincial Attorney General's Department but accountable also to the Ministry of Education (which provides indirect funding through release-time for teachers, some transportation assistance, etc.).

In each of these case studies the structural context of the programs and the level of involvement of the various evaluation negotiators was central to the subsequent use of the evaluation findings. Both programs were situated in structural contexts where they have had to "prove themselves" to those who had the power to set policy and provide funds for their ongoing operation. The decision to undertake an evaluation came from within the programs themselves (not from outside/above), and reflect an effort by members of the research setting (who were simultaneously the evaluation funders) to document the scope of their present impact and provide direction for future development.

**The JFSA Evaluation: Conclusions and Implications**

In my analysis of the JFSA evaluation it became evident that the negotiations regarding the agency assessment were only one of two distinct sets of negotiations going on at the same time. Preceding our hiring as consultants,
the Board and the Executive Director had been engaged in negotiations regarding the scope and limits of her power in administering the agency. These negotiations, while relatively covert, were ongoing and serious for all parties concerned. The conflict between the Board and director was never resolved; it just ended with payment of an out-of-court settlement. However, the negotiations regarding evaluation of the agency had a different focus and set of issues. We refused to become involved in the original conflict, defining it as outside the legitimate boundaries of our authority and competence. Our unit of analysis was the agency as an organizational whole, and our concern was with its operation in the larger Jewish community of the Lower Mainland and in the context of other JFSA's in cities with similar size Jewish populations.

The power struggle obviously reduced the level of involvement of the director in the evaluation process, although she permitted consultation and occasionally participated in the data gathering process. Members of the staff, too, were distracted by the ongoing conflict between the director and Board, and were unused to having any input into policy decisions in the agency. They were, however, pleased to be part of the evaluation process and were active participants in the research activities. The nearly twenty members of the Board ran the gamut from total disinterest, to a core group who participated actively in planning and discussion of the evaluation study. This culminated in the pattern of ownership exhibited by a subsequent Board President who guided and organized the intense effort to achieve increased agency funding and improved staffing for the development of expanded programming.

Reflecting on the situation at a later date, I noted that both sets of negotiations, the conflict between the Board and Director and the more
cooperative agency assessment process, have never been concluded since there is still no long-term planning process or management policy which has been developed within the agency. I suggested that some or all of this legacy might be the result of several factors in the structural context of the agency which could never be remedied by a single assessment study. These factors included: the low level of financial and political support for local needs among powerful members of the Jewish community; the low levels of involvement in policymaking regarding local social, cultural and economic needs among a cross-section of community members; the limited power and relatively low status of the JFSA Board members as perceived by the community; and the basic inability of these Board members to define their own power and allocate clear responsibilities to their director and staff.

This overall pattern of connection between the structural and negotiation contexts of the evaluation of this agency underlines the extent to which program evaluation work becomes inextricably linked with the other issues and negotiations which are ongoing in an organization. In addition, the case study offered an example of the way in which other negotiations can limit the levels of involvement of negotiators in the evaluation process. And, finally, it presented a case where the research model was never at issue, the research report was well received and utilized in subsequent program planning. But the essential power over long-term effectiveness of this agency lay in the hands of those negotiators who had always remained outside our negotiations.

In this case study again there are implications regarding the appropriate unit of analysis and the problem of defining the appropriate organizational boundaries for evaluation research. It has been argued that when evaluation studies are national in scope and employ a single research model
they run the risk of overgeneralization and a lack of accuracy and validity of findings. However, in this instance, where the study unit was limited in size and scope there are two different problems. First, there is the objection that the research results may be idiosyncratic, perhaps reflecting only the biases of the evaluators. Second, there is the difficulty that evaluation of a single agency or institution in a community is seldom sufficient to have an impact on the structural context of the community as a whole. These problems may also emerge in the study of large scale national programs where each sub-unit is evaluated using a different research strategy, model and methodology. There is no final answer to these questions regarding the appropriate unit of analysis, however it has emerged as a significant aspect of the research model selected for use in the evaluation of large and small-scale programs.

LSS Schools Project Evaluation: Conclusions and Implications

There were several characteristics of the evaluation of the LSS Schools Program which contrasted with the preceding cases. This last case study offered an alternative set of insights into the questions of evaluation placement, research model and utilization of results. To begin with, I was hired as a resource person to consult with all members of the Schools Program, not as an "outside evaluator". Second, the entire evaluation process and subsequent planning and implementation of changes in the program entailed a collaborative relationship among the staff, summer research student and consultant. Finally, the evaluation report represented the last phase of our work together and was written collectively after a) thorough discussion and analysis of research findings and, b) changes in the program policy and activities had been implemented in the day-to-day operation of the program.
Among the four evaluation studies the transactions in this situation were the most visible and most organizational issues were negotiable. The one significant issue over which there was little or no overt negotiation was the dissatisfaction in the relationship between the two main staff members and the director. This critical problem had developed out of differential approaches to the internal balance of power and was reflected in concerns over the unclear division of labour, ideological conflicts, levels of competence and working styles. It appeared that the tasks of the evaluation process resulted in a redistribution of work responsibilities and some efforts to implement a more equalitarian distribution of power within the day-to-day operations of the program. This permitted the director and staff members to defer/deflect a possible confrontation over the question of power and responsibility.

Despite this arrangement where the balance of power within the program was kept beyond the boundaries of the negotiation context, the actual evaluation process was given full attention by the staff and director. They were involved in all the basic areas of negotiation, and ownership of the evaluation process and product was located within the Schools Program. This pattern of ownership may be interpreted as the source, cause or context for the relatively quick and successful series of organizational changes implemented within the Schools Program during the two year period of the evaluation study. Findings in this study suggested that where there is a relatively high level of involvement of negotiators across all the basic areas of negotiation, there may be a greater potential for the evaluation results to be rapidly integrated as part of a process of organizational change.

Given the success of this evaluation process within the Schools Program, we must note that there were many factors in the structural context of the
Legal Services Society which tended to undermine the achievement of substantial long term changes in program operation. Essentially, the obstacles to permanent changes in the management and scope of the Schools Program lay in the strongly hierarchical and instrumental approach which characterized the upper level of management of the LSS and its Board of Directors. Their emphasis on control over subordinate levels in the organization and a resource shift to the more traditional aspects of Legal Aid ran counter to the procedures and objectives of the Schools Program. Unfortunately the pattern of minimal planning and weak management characteristic of Legal Information Services had tended to increase the vulnerability of the Schools Program as well as the other sub-units of the overall LIS component. The policy of control and/or "no response" at upper management levels had prevented all efforts to introduce democratic/equalitarian forms of decision-making. The strike at LSS in Spring 1981, was the most overt indication of a wide range of problems in that structural context.

One consequence of these structural pressures at the LSS was an externally imposed resolution of the hierarchical/equalitarian controversy over the balance of power within the Schools Program. That is, the Director of LIS had replaced the previous Director of the Schools Program (who resigned to return to university for an advanced degree) with a person who shared the managerial ideology and practices of the upper levels of management and the majority of the Board of Directors of LSS.

The main conclusions of this study point to the critical significance of examining: a) the internal balance of power and responsibility within the research setting and b) the appropriate unit of analysis for initiating and reinforcing a process of long-term organizational change. In analyzing both the JFSA assessment and the Schools Program evaluation it has become
evident that unresolved conflicts among major actors (board, director, staff) regarding the balance of power and responsibility within the research setting may set constraints on the pattern of involvement in the evaluation process. However, even in a setting where the evaluation process entailed a high level of involvement and satisfactory collaboration in achieving organizational changes, the chief obstacles to maintaining those changes emerged from the overall structural context of the program in question. In some research settings the mere fact of being involved in any form of evaluation study may either emphasize or deflect the ongoing conflicts over power and responsibility. Some program contexts require that the evaluator specify clearly whether that ongoing conflict is to be included or excluded from direct consideration in the evaluation process. This is an aspect of the structural context of programs that may become as significant as the powerful negotiators who remain outside the boundaries of the evaluation process.

Summary of Four Evaluation Case Studies

The following schematic summary (Figure 4) provides an overview of the four case studies in terms of the main components of their structural context and the negotiated arrangements and outcomes in each of the program evaluation studies. The summary was set up to simplify some of the comparisons across case studies which I did not include in the text. However, the design of the summary does not permit me to show the detailed relationships between aspects of structural context and the evaluation negotiations I have described in the preceding individual analyses of the cases.
### FIGURE 4 SUMMARY OF FOUR EVALUATION CASE STUDIES

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Shift from urban renewal to residential &amp; neighbourhood rehabilitation of inner city ethnic neighbourhood; pilot project funded by 3 levels of government; inclusion of community group (SPOTA) on management committee (SRC); major control and allocation of resources in government hands; consensus model of decision making despite barriers to participation, relatively high proportion of citizen participation in grant/loan program; citizens subsidize heavy funding allocation to underground services not rehabilitation component.</td>
<td>Separation of neighbourhood improvements &amp; residential rehabilitation; funding from 3 levels of government; national programs developed to serve urban &amp; rural areas, require citizen participation on local NIP committees involves complex site eligibility criteria to be met; specific funding allocations by province; funded for period of 4 years (1974-78); proposed to be implemented in at least 400 locations, but may not have reached that total before termination; financial eligibility criteria for RRAP participation.</td>
<td>Private social service agency serving Jewish community in lower Mainland and throughout B.C.; jointly funded by United Way and Jewish community fund and council; in existence more than 50 years carrying out service needs to immigrants, elderly and others needing assistance; not meeting basic community needs because underfunded &amp; understaffed; low status in community along with other local educational/cultural/service institutions.</td>
<td>Part of Legal Information Services; recent organizational &quot;marriage&quot; of Legal Services Commission &amp; Legal Aid; Curriculum development, training, publications &amp; audiovisual products for teachers &amp; students at secondary schools throughout B.C.; recent rapid growth in program demands &amp; responsibilities; provides funding for independent development of curriculum &amp; A/V materials; own system of sales &amp; distribution, also through Ministry of Education system.</td>
<td></td>
</tr>
<tr>
<td>Major Actors</td>
<td>SRP Evaluation</td>
<td>NIP/RRAP Evaluation</td>
<td>JFSA Assessment</td>
<td>LSS Schools Program Evaluation</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------</td>
<td>---------------------</td>
<td>----------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Multiple, four distinct interest groups, local community members, service workers, and 9 person study team; contributed to study complexity and delays.</td>
<td>NRID Director representing CMHC, National Study Coordinator and 5 person study team (split between Ptarmigan and CPAC). Unrepresented but powerful members of CMHC management group Ottawa</td>
<td>Board of 20 members (5 active), director and 3 person staff, plus study team with 2 members. Unrepresented but powerful members of community fundraisers/elite.</td>
<td>Program director, 2 staff, a secretary and summer student, assisted by the evaluator as consultant. Unrepresented but powerful LSS board &amp; management group, and LIS director.</td>
<td></td>
</tr>
<tr>
<td>Organizational Placement of Evaluation</td>
<td>External; accomplished after Stage 1 Evaluation; most project components completed; some unfinished project business; choice of evaluator determined by SRC.</td>
<td>External; determined by CMHC Head Office management group; coordinated by Study Director in Toronto &amp; &quot;farmed out&quot; to research teams in seven provinces.</td>
<td>Generally external to agency, commissioned by the board, some research done by director and staff members; Board President took &quot;ownership&quot; of the evaluation process/product.</td>
<td>Internal to Schools Program, Commissioned by director and staff who did major research tasks, some analysis, decided on program changes &amp; owned evaluation process/product, evaluator as consultant.</td>
</tr>
<tr>
<td>Research Strategy</td>
<td>Consensual approach at beginning; economic analysis encouraged elements of contentious approach re actual project allocations.</td>
<td>Consensual approach with reference to goals/objectives of NIP program; contentious and paradigm challenging with respect to strategy and model employed by national study director.</td>
<td>Consensual approach, basically a managerial review with extensive comparison to other agencies; findings not contentious, but politics of evaluators (perhaps).</td>
<td>Elements of all three approaches, depending upon program component under review and proposed organizational changes.</td>
</tr>
<tr>
<td></td>
<td>SRP Evaluation</td>
<td>NIP/RRAP Evaluation</td>
<td>JFSA Assessment</td>
<td>LSS Schools Program Evaluation</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------------------------</td>
<td>--------------------------------------</td>
<td>----------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td><strong>Model and Methods</strong></td>
<td>Wide range (see appendix 1) use of both models in overall research design; emphasis on agreement between numerical and observational results.</td>
<td>More limited than SRP, but involved use of BC program records analysis of budgets, observations &amp; interviews at meetings &amp; survey of 54 actors involved in various aspects in four case study sites.</td>
<td>Elements of both models used; file data (collected by staff to protect names) 50 interviews with key informants; financial &amp; service data on 16 other agencies; budgets.</td>
<td>Elements of both: sales/distribution analysis, survey of users, formal/informal interviews, analysis of budgets &amp; funding policies, tasks and division of labour.</td>
</tr>
<tr>
<td><strong>Research Report</strong></td>
<td>Two stage construction process; confrontation with SRC over first draft; minimal difficulty in division of labour within study team; draft &amp; final report very late.</td>
<td>Source of substantial study team conflict/disagreement; but ultimate completion; focus of conflict with CMHC Head Office (re anonymity); on time, constructed according to study team design.</td>
<td>No problems in division of labour and report construction; results summarized and reported in annual general meeting; some obstacles to distribution of written report.</td>
<td>Collaborative process of production; assembled in order to describe &amp; document changes in process/structure of program.</td>
</tr>
<tr>
<td><strong>Research Utilization</strong></td>
<td>Delivered into &quot;policy limbo&quot;; no plans or procedures for subsequent use; no connections to Ottawa policy planners.</td>
<td>Some input of results at Action Conferences; however program terminated and reduced funding shifted to provincial control.</td>
<td>Used extensively by board president in successful campaign to increase agency funding seen as a policy &quot;straight-jacket&quot; five years later.</td>
<td>Results used immediately to make organizational changes; evaluation report useful in describing/explaining program significance.</td>
</tr>
</tbody>
</table>
6.3 Preliminary Formulations: Models of Negotiated Knowledge Production and Use

Background to the Models

Before delineating the proposed models of knowledge production and use, I want to indicate something of the "grounded theory" approach which led to their formulation. The original data collection which was done in each of the four case study settings was not totally controlled by the immediate needs of producing a program evaluation document. As the research director/consultant in each case study situation I was guided implicitly by an ongoing concern with the possible relationships between the work of social research, the policy intervention theories which are supposed to provide the basis for the development of reform programs, and the actual operation of such programs in the field. In beginning the dissertation I had a partial framework gained from extensive work in social reform programs and a range of policy related research. I was willing to make a commitment to the negotiated order paradigm and the theoretical work done by Rein and Wiseman since they "fit" with my own experiences and theoretical preferences for approaches which permit the generation of theory from data systematically obtained in those settings to which the theory will have direct applicability. The sample of case studies and the data selected for description and analysis in this dissertation were shaped by my emerging interest in making substantive generalizations about how policy research is constructed and the extent to which it may or may not be used in specific programs and policy contexts. They were selected in order to provide a set of comparative situations from which I might begin to integrate conclusions, concepts and hypotheses into preliminary formulations of the expanded set of models. I view these models as an initial
stage in generating substantive theory regarding program evaluation as a form of policy research (Glaser and Strauss 1967:118).

In the comparative analysis of the case studies, summarized in Figure 4, it is possible to trace two different types of substantive relationships in the evaluation process. First, reading down the individual columns one can follow through each case study, starting with those critical elements in the structural context of the program which set limits and/or facilitated a pattern of knowledge production and use in the negotiation contexts of the program evaluation process. In addition, it is possible to read across categories and compare significant components of the structural context of each evaluation case study as well as the range of "solutions" to the essential theoretical/practical issues of the program evaluation as reflected in each of the categories defined as basic areas of negotiation.

The conclusions drawn from comparing these cases have generated: (a) a limited series of substantive generalizations which could go through a verification process in an additional, much larger sample of program evaluation case studies, and (b) implications for the construction of an expanded set of research models specifically applicable to program evaluation settings. These generalizations would include such statements as:

- Where powerful actors are unrepresented in evaluation negotiations, there is a reduced possibility that proposed long-term program changes will be implemented.

- The unit of analysis most appropriate for initiating long-term organizational change is likely to differ from the unit of analysis defined by the major actors who control the funding of program evaluation work.

- The internal balance of power and responsibility within the research setting (program to be evaluated) is likely to have a significant impact on all the basic areas of negotiation in the evaluation process.
Preliminary assessment of the major components of the structural context of the program (social, political, economic, and organizational) will tend to reduce obstacles and delays in constructing and presenting the evaluation document.

Early planning for research utilization and the development of contacts with significant intermediaries/power brokers will increase the likelihood that the evaluation document is read and findings used in some specifiable way.

There were several implications for the formulation of the models. Detailed examination of the case studies suggested that in each evaluation study the decisions made and implemented in one specific negotiation context (e.g., organizational placement, research strategy) tended to be related to an overall patterning of negotiations in the other basic areas of negotiation. For example, the organizational placement of both the SRP and NIP/RRAP evaluations contributed directly to the lack of research utilization, despite the good rapport developed with respondents and program staff/managers within the local communities where the evaluation research was conducted. Looking specifically at research strategy as a basic area of negotiation, it became evident that what might have appeared to be a consensual approach could not necessarily continue to be consensual as the evaluation process continued into further negotiation contexts. That is, the model and methods preferred by any one or more of the major actors might force a reinterpretation and relabelling of an evaluation team's research strategy, as well as their model, method and findings. The ideological preferences of the major actors generally have an impact on the construction and utilization of the evaluation research report. For example, the use of "straightforward" economic data and conventional financial analysis is no guarantee of the acceptability of evaluation findings when major actors have pre-determined prejudices regarding the program target group (SRP results which documented a significant level of financial input from low income Chinese residents). Under some
circumstances, a pattern of ambiguity and/or conflict that characterizes the initial relations between major actors and evaluators may reflect essential differences in research strategies, subsequently affecting all the remaining areas of negotiation in the evaluation process (the NIP/RRAP evaluation). Finally, other sets of organizational negotiations may enhance or detract from the evaluation process; for evaluation negotiations often become contingent on the existing degree of cooperation or conflict which characterizes the immediate structural context of the program under review (evident in both the JFSA and the LSS Schools Program evaluations).

These are some of the salient features of the case studies which suggested the need for a further process of abstracting and typifying the basic areas of negotiation into different approaches to the question of knowledge production and use. Initially this involved consideration of the extent to which new models could incorporate systematic recognition of the characteristics of the structural context of the programs being evaluated. Findings from the case studies and other theoretical sources suggested that such questions as the representation of major actors, the organizational placement of research and the use of the evaluation findings could not be clearly interpreted without an understanding of the structural arrangements in which the program was planned and implemented. Further, preferences regarding research strategies, models and methods were evidently shaped by several influences in the structural context and professional training of policymakers, program managers and evaluation researchers.

In developing the preliminary forms of these models of negotiated knowledge production and use I had to rely on several other empirical studies of program evaluation and policy research (especially those presented in Abt 1976; Cronbach 1980; Guttentag and Struening 1975 a, b; Rowbottom 1977;
and Weiss 1977) as well as the conclusions and implications drawn from the four case studies presented in the preceding chapters. Analysis of these several sources has suggested four preliminary models of negotiated knowledge production and use. These models include elements of the positivist and interpretive models discussed in Chapter 2; however they are designed to take into account aspects of the structural context of program planning and implementation and focus directly on the six basic areas of negotiation in the policy research process.

This expanded repertoire of models of negotiated knowledge production and use have been labelled experimental, managerial, collaborative, and transformative. The labels have been developed partly as quick descriptions, but mainly because those terms represent the particular area of negotiation which tends to dominate other aspects of the particular research model. I have constructed the models in terms of their direct applicability to the work of program evaluation, and one direction for further research might involve exploring their use in various social impact assessments and planning studies for future organizational change in selected reform programs. In the following discussions of each model I have begun by describing the general characteristics of each of the basic areas of negotiation in the evaluation process and suggest selected characteristics of the structural context in which such a model might be followed. Finally, I present some brief examples from case studies or other sources which indicate the circumstances under which aspects of the model were used in specific empirical settings. It should be understood that all program evaluations are "mixed models" and include some elements from most of the ideal types I am describing here.
The Experimental Model

Generally in this experimental model of knowledge production and use, the program evaluator takes the familiar consensual approach to overall research strategy, sharing with the major actors (usually limited to representatives of the funding agency and management of the program) a positivist model of evaluation research and reliance on objective measures of program impact. All the evaluation research is designed to be accomplished by evaluators whose work base is external to the program under scrutiny. High technical standards are important for the evaluation research and close adherence to experimental design are key goals for program managers and staff. Therefore, the resulting research report is constructed in a negotiation context where conflict is minimized and objective, numerical evaluation findings are given great credence. The final evaluation document is presented to the funding agency for interpretation and future use in the policy process. There is minimal evaluator input into subsequent phases of policy development and program modification.

This approach might be suitable, for example, in a school system (or series of systems in smaller communities) where there has already been some "mapping of the territory", and substantial observational work completed. In those communities where there is a relatively high degree of cultural homogeneity and there are agreed-upon policy intervention strategies that representatives of the school system and community wish to assess and compare, this might be the most appropriate model for program evaluation research.

There are other structural contexts in which such a research model might be used as an adjunct to other types of organizational and observational research studies. Examples would include such total institutions as prisons and mental hospitals where different theories are supposed to be implemented
in the treatment process. Under these conditions, the experimental model might be used to guide evaluation research regarding the impact of such treatment on selected groups in the inmate/patient population. However, without the other type of studies which take into account the structure and process aspects of treatment programs the experimental model tends to be an inadequate and sometimes misleading tool for evaluation research.

There was no clearly developed example among the case studies of this model of knowledge production and use. However, it did appear that aspects of this model were implicit in the overall evaluation research plan designed by the national coordinator of the NIP/RRAP evaluation. There was no question of randomization in the design of the program itself, but within the evaluation design there was an emphasis on control and pre-determination of all variables to be included in the survey instrument. Members of the British Columbia study team never saw the completed form of the national plan, and given the enormous variation in program development and implementation in each province, we considered that such a highly structured model and methods had limited relevance to the policy questions put forward by the management group at CMHC. However, there have been several other types of structural contexts in which the experimental model has been the preferred approach, especially in education (Cronbach 1980:274) and health care delivery (Rein 1976:187). Specific experiments in the impact of the guaranteed annual income (Gilbert 1980) and in the home care of schizophrenics (Franklin and Thrasher 1976:52) offer examples of the range of structural context in which this model has provided the basic framework for both program planning and evaluation studies of program effects.
The Managerial Model

In the second model of knowledge production and use the major impetus for evaluation work comes from the senior level of the organization, and the research is defined as part of the overall cycle of planning, decision-making, implementing and evaluating. The evaluation report is likely to be one of several documents used in planning for the next cycle of organizational change. Therefore, this is the model frequently employed in managerial reviews of ongoing programs or program components where the evaluation research is done by "in-house" researchers and the major actors are all members of one organizational context (e.g., Ministry of Economic Development, Central Mortgage and Housing).

In such structural contexts the evaluator(s) employ a consensual research strategy, where consensus implies acceptance of the theory and practice of "good management" and is the main focus of program assessment. There is an explicit emphasis on the discovery of the day-to-day facts of program operation, careful recording of individual accomplishments and failures, and a steady reliance on whatever research models and methods have been followed by staff in assembling the data required by program monitoring procedures. This does not always mean quantification of everything under scrutiny, but it does mean accepting the categories and labels given to data in the ongoing process of program management, e.g., substandard housing, unemployment rates, schizophrenic patients, persons ineligible for program benefits, etc. There is an emphasis on statistical or observational conclusions which are not "counter-intuitive" to the accepted wisdom of the Ministry/agency, and the resulting evaluation report may frequently be used as grounds for firing personnel, eliminating a project/department, or alternatively as a means to cover up problems in program delivery which have been defined deliberately as lying outside the parameters of the management review.
This model of evaluation research has tended to be the standard approach employed by the Government Accounting Office in the United States and the approach preferred by the Treasury Board of Canada and the Comptroller General (1981). The president of the Treasury Board views the program evaluation "function" being established in federal departments and agencies as "an essential part of the government's initiative to improve management practices and controls" (1981:2).

Generally this was the model we used in doing both the NIP/RRAP and JFSA evaluations, and it was also the core of the SRP Stage II evaluation. In retrospect, it seems unfortunate that such a simple model of program monitoring and evaluation was not built into the initial planning phase of the SRP and several earlier housing programs funded by the senior levels of government and discussed by Dennis and Fish (1972). However, up until recently there was no required procedure or agreed-upon model for evaluation built into the implementation phase of large-scale government-funded programs. Despite strong criticisms of this model by community advocates representing program clients/participants (often appropriate criticisms deriving from another research perspective), the absence of any type of evaluative information on program process and/or impact may be a more powerful obstacle to ensuring program accountability and increased responsiveness to community needs.

I would not downplay the significance of the assumptions regarding the status quo built into this model, but it can serve several unanticipated purposes in the process of change. A managerial review may become the first step in creating greater awareness of local needs among program clients/participants and a basis for demanding an alternative evaluation which employs a research strategy which takes local needs and priorities into account (Siegel and Doty 1978).
The Collaborative Model

Generally, the collaborative model has been used in organizations attempting to make gradual, long-term changes in their structure and day-to-day operation. Frequently, these are programs which have reached some type of organizational crisis point in terms of work overload/employee turnover/rapidly shifting demands for program service; they are looking for a way to evaluate their existing situation and make systematic improvements. In this model there is an attempt to involve as many actors or representatives of interested parties as the evaluation process can effectively include. The evaluation is located both externally to the program (often through use of specific consultants) and simultaneously as the "in-house" process of self-analysis and internal research. Given that there are conflicting theories of intervention represented in most reform programs, the evaluation negotiations frequently centre on developing a research strategy that is minimally acceptable to all parties. Then the evaluation researchers tend to make use of an array of models and methods which will be approved and understood by different components of the program's staff and service users. The final form of the evaluation report is developed collectively as a documentary review of program changes in process, including minority policy positions and reports of program inadequacies and failures. Although some form of report goes into "document time", most research results are utilized immediately by various parties to the evaluation through a process of continual discussion and feedback. Since this model tends to focus on a form of collaborative effort, small-scale conflicts can usually be managed within the negotiation process. However, significant unresolved differences in intervention theory and major inequities in power and resources among the potential partners can often produce serious obstacles to the achievement of long-term
organizational changes in program policy and/or operation. Effective use of this model of evaluation research presupposes a careful determination of the appropriate unit of analysis in the study, widespread involvement in the planning and implementation of organizational changes suggested by the research findings, and an ongoing strategy for conflict resolution at all levels of the organization undergoing change.

This third model of knowledge production and use which characterized the evaluation process in the LSS Schools Program has a relatively long history of use in several industrial and health/social service settings. As documented in Rowbottom's (1977) text on "social analysis", which is defined as a collaborative method of gaining usable scientific knowledge of social institutions, this model was originally employed by Elliot Jacques in the late 1940's (Rowbottom 1977:v). However in recent decades it has been employed in the total reorganization of the British National Health Service and in local authority Social Service Departments across that country. The model has also been used for work on hospital organization (Rowbottom et al 1973).

The Transformative Model

The fourth model of knowledge production and use may be the least common approach to policy research presently in use. In this model program evaluators seek active support and involvement of those actors least likely to be represented in the other models, i.e., the program participants and service users/consumers. The evaluation is specifically designed to be external to the program and at some conceptual distance from the theories and assumptions of program funders and managers. The research strategy tends to range from the contentious approach to a paradigm-challenging stance which calls into
question program policy, funding, structure and operation. The choice of research models and methods may vary, but where the analysis relies on a political economy or Marxist approach there may be significant reliance on positivist models and an emphasis on objective methodologies. The methodological details may depend to some degree on the parties negotiating and the constituency to be addressed in the evaluation document. Frequently the final form of the evaluation report is developed by the evaluator(s) and other community representatives as a document to provoke critical discussion of program policy, operation and change. It is structured to emphasize the gaps between program theory/goals and the limited objectives of day-to-day program services. Program evaluators and community representatives often plan how research results may be used to educate program clients/consumers and to refute the findings of program evaluations which have been done using a managerial or experimental model of knowledge production and use.

The area of health care needs and service delivery appears to be one of the main structural contexts in which a variety of paradigm-challenging evaluations have been done. These studies have ranged in scale from evaluation of a single project to an assessment of how social, political, economic and ideological factors have shaped the emergence of a new health care reform organization. The evaluation of a single alternative health care structure was Kleiber and Light's (1978) study of the Vancouver Women's Health Care Collective. They used a form of collaborative model in order to assist the collective in scrutinizing their day-to-day operation and longer term objectives, while simultaneously documenting the resulting changes which collective members decided they must make in their division of labour and coordination or tasks and responsibilities. Health and Welfare Canada funded the documentary process and the salaries of the two researchers/consultants. However,
the basic approach of the Health Collective challenged traditional Canadian health care delivery paradigms and the Kleiber and Light study used a research strategy which fit with the collective's values and ideology. Thus the researchers' approach, by definition, was viewed as paradigm challenging by selected members of the technical advisory board of Health and Welfare Canada who were overseeing the research.

At the next level of scale and complexity, there was the Chu and Trotter (1974, discussed in Siegel and Doty 1978) report on the Community Health Center program. This evaluation study produced for the Center for the Study of Responsive Law (Nader's Raiders), is described as a consumer evaluation of the program, and includes several paradigm-challenging recommendations for program change which are not supported by the overall research strategy employed in developing the research findings.

The final study referred to above is Hessler and Beavert's (1982) summary assessment of a longitudinal study of neighbourhood health centers (NHCs) for the poor in the United States. They present an analysis of the political economy of change within reform organizations and seek to explain the shift in the role of poor people participating in health care decision-making from that of program developer and change agent to the role of program restrictor. This prospective study of NHCs between 1965 and 1977 (see also Hessler and New 1972; New and Hessler 1973; Hessler New and May 1979) represents an evaluation of changes in community participation, health care ideologies, and funding criteria which have taken place during more than a decade of structural changes in "the political and economic circumstances that gave rise to the NHC as a reform organization with the strategy of citizen control" (Hessler and Beavert 1982:253). This study is paradigm-challenging in that it does not take the present arrangement in NHCs as a
given, but has traced the large scale shifts in the structural context of the NHCs, the opportunities and constraints they presented during the 1960's and the opportunities and constraints they provided during the changed political and economic climate of the late 1970's.

6.4 Possible Directions for Future Research

These four models of knowledge production and use suggest several directions for future research. One possibility is their use in other types of policy research settings such as social and environmental impact statements which are required by several government departments/Ministries. However, since I am now most familiar with the structural and negotiation contexts of the work settings in which program evaluations are commissioned, constructed and received, it is that part of the policy process that I would study in greater scope and detail.

Events in the past decade have given program evaluation a new and expanded function within the Federal levels of the Canadian Government. During 1981 the Treasury Board of Canada and Comptroller General funded the printing of two companion volumes on program evaluation. They are addressed to deputy heads of departments and agencies and to other officials responsible for program evaluation, and are intended for use by other levels of government and other organizations interested in the evaluation of government programs.

The first volume, *Guide on the Program Evaluation Function*, "describes the systems and procedures departments and agencies in the federal government are to have in place to ensure a useful and relevant program evaluation function, and explains the general approaches and principles to be used in
carrying out evaluations of programs" (1981:1). The companion volume, *Principles for the Evaluation of Programs by Federal Departments and Agencies*, deals with the actual process of program evaluation including principles of conducting, reporting, and standards for maintaining the quality of evaluation assessment studies and evaluation studies. In the closing paragraph of the Foreward, the Comptroller General remarks that "the principles set down here should not be looked upon as substitutes for common sense, or for the ability to select the correct approach to a program evaluation. They reflect, indeed, the view that evaluation, like any other management function, is an exercise in professional judgement."

From my perspective these companion volumes offer an ideal place from which to initiate an extensive, long-term description and analysis of a sample of evaluation assessment studies and evaluation studies commissioned by the deputy heads of several departments. The mandate of these volumes is that program evaluation "should be an integral part of the management review and monitoring function in departments and agencies, providing input into planning and budgeting" (1981:vol. 1:11). This mandate would offer an opportunity to study the structural context and research models employed in doing departmental evaluations and trace the subsequent uses of those evaluation documents in the policy planning and budgeting cycles.

Using the framework developed in Chapter 3 would offer a unique theoretical approach through which the evaluation negotiations and consequences could be linked with the structural context of program planning and implementation. In addition, the four models of knowledge production and use might be tested and expanded through my extensive, long-term analysis of the various evaluation studies initiated in response to Treasury Board's new guidelines and requirements. In particular, I would be interested in observing and
analyzing those structural arrangements which facilitated overall involvement in the evaluation process, obstacles to the implementation of recommended changes, and the types of research procedures and reports which were likely to evoke cooperation/conflict in differing governmental departments and agencies.

In studying a sample of the existing and future evaluation studies accomplished under the mandate of Treasury Board's guidelines, it would be essential to begin by focusing on a single department in order to document the linkages between the structural context and the various models of program evaluation produced and used within that context. There are several stages of preliminary work which would be necessary in order to arrive at a research proposal with manageable dimensions. These tasks would include first of all careful analysis of the purpose and content of the guidelines and interviews with those who produced the two volumes for Treasury Board. Second, there would have to be a series of negotiations with selected departments/agencies in order to discover one or more which might be willing and interested in participating in a study of this nature.

The remaining research tasks would involve the relationship with the major actors within the department or agency selected as the site for this work. This would entail a preliminary assessment of the structural context and existing evaluation strategies in use within the department. In addition, the major actors at several levels in the organization would have to be convinced to commit themselves (and those whom they represent) to a mutually agreeable level of involvement in this long-term, overall study of the department cycle of planning, decision-making, implementation and evaluation. Although emphasis would be on the evaluation models, such an analysis could not be completed without data on the other phases of the policy planning process.
Description and analysis of the models of knowledge production and use developed within the structural context of programs funded by a single government department would provide an understanding of the extent to which certain research models are likely to meet specific policy planning needs.
1. I had extensive familiarity with several aspects of the Health Collective's development and operation. As field staff person for the CYC (Company of Young Canadians, Canada's domestic Peace Corps), I supervised two key members of the Collective who worked as organizers on CYC salaries, and was simultaneously a member of the Technical Advisory Committee established by Health and Welfare.
BIBLIOGRAPHY
BIBLIOGRAPHY

Abramson, M.A.  

Abt, Clark C., ed.  

Alinsky, Saul D.  

Altshuler, Alan  

Arnstein, Sherry R.  

Axworthy, Lloyd, ed.  

Baker, K.  

Barnard, Peter and Associates  

Bean, N.  

Becker, Howard S.  

Bell, Colin  
Bell, Larry and Richard Moore

Benson, J. and R. Day

Bernstein, P. and L. Bowers

Blackburn, Robin, ed.

Bogdan, Robert and Margret Ksander

Boruch, R.F., and H.W. Riecken, eds.

Bowles, Roy Tyler

Britain, Gerald M.

Britain, Gerald M., and Ronald Cohen, eds.

Brooks, Michael P.

Burrell, G. and Gareth Morgan

Caplan, Nathan

Caputo, David C.

Caro, F.G.

Carter, Novia and Brian Wharf
1973 Evaluating Social Development Programs. Ottawa: Canadian Council on Social Development.

Cassell, Joan

Central Mortgage and Housing Corporation

Central Mortgage and Housing Corporation, Neighbourhood and Residential Improvement Division

Chapin, F. Stuart, Jr.

Cicourel, Aaron V.

Clairmont, Donald H. and Dennis W. Magill

Cole, Richard L.

Coleman, J.S.

Colvard, Richard
Community Planning Association of Canada (British Columbia) and Ptarmigan Planning Associates

Crenna, David

Cronbach, Lee J. and Associates

Crowfoot, Jim
1977 Organizational Elements to be Considered in Worker Democratization. Diagram for Democratic Management Seminar, School of Ecology, University of Michigan.

Curtis, James E. and John W. Petras, eds.

Dalton, Melville

Daniels, Arlene Kaplan

Datta, Lois-Ellen

Day, R. and J. Day

Deaton, Rick

Dennis, Michael and Susan Fish
Denzin, Norman K.

Deutscher, Irwin

Dillman, Caroline M.

Douglas, Jack, ed.

Draper, James A., ed.

Drucker, Peter

Eaton, Joseph W.

Enoch, Robert and Alberta Potter Levitan

Etzioni, Amitai

Etzioni, Amitai, ed.

Fay, Brian

Fetterman, David M.

Finsterbusch, Kurt and Mary R. Hamilton
Finsterbusch, Kurt and C.P. Wolfe, eds.

Franklin, Jack L. and Jean H. Thrasher

Fraser, Graham

Freeman, Howard E. and Clarence C. Sherwood

Fried, Marc

Frieden, Bernard J. and Robert Morris, eds.

Galliher, John F.

Gans, Herbert

Geertz, Clifford

Giddens, Anthony

Gilbert, Sid

Gilmore, John S. and Mary Duff

Glaser, Barney G. and Anselm Strauss
Glass, Gene V., ed.

Goldberg, Michael
1976 Strathcona Community Study Phase I: Social Science Resources.
   Vancouver: Strathcona Community Resources Board.
1977 Strathcona Community Study Phase II: Residents of the Community.
   Vancouver: Strathcona Community Resources Board.

Goldberg, Michael and Alberta Potter Levitan
1977 Vancouver Jewish Family Service Agency: An Assessment. Vancouver:
   Ptarmigan Planning Associates.

Goldner, Fred H.
   245-266.

Gouldner, Alvin W.

Gouldner, Alvin W. and S.M. Miller

Greenberg, E.S.

Gutstein, Donald
1975 Vancouver Ltd. Toronto: James Lorimer & Co.

Guttentag, Marcia, Lorelei R. Brush, Alice Ross Gold, Marnie W. Mueller, Sheila Tobias and Marnie Goldstein White

Guttentag, M. and E.L. Struening, eds.

Hall, Budd L.
   8(2):24-36.

Hammond, Philip E., ed.
Handleman, Don and Elliott Leyton  

Harp, John and John R. Hofley, eds.  

Heap, James L.  

Heap, James L. and Phillip A. Roth  

Herbst, Phillip G.  

Hessler, Richard M. and Carolyn Sue Beavert  

Hessler, Richard M. and Peter Kong-Ming New  

Hessler, Richard M., Peter Kong-Ming New and J. Thomas May  

Holleb, Gordon and W.H. Abrams  

Horowitz, Irving Louis, ed.  

Horton, John  

House, Ernest R.  
Hurwitz, R. Laura  

Hymes, Dell, ed.  

Jacobs, Sue-Ellen  

Johnson, John M.  

Kahn, Si  

Kanter, Rosabeth Moss  

Kaplan, Abraham  


Kaplan, Michael  

Kleiber, Nancy and Linda Light  

Knorr, Karin D.  

Kuhn, Thomas S.  

Laue, James H.  
Layard, Richard, ed.

Lazersfeld, Paul F.

Lazersfeld, Paul F., Jeffrey G. Reitz and Ann K. Passannella

Lee, Dorothy

Legal Services Society Schools Program

Lehman, Edward W.

Levitan, Alberta Potter and Laurie Hurwitz

Light, R.J. and P.V. Smith

Lindblom, Charles E. and David K. Cohen

Lynd, Robert S.

Mackie, Marlene, ed.
1980 Sociology's Relations with the Community. Calgary: University of Calgary Department of Sociology Colloquium Series.

Madge, John

Mahon, Rianne
Maines, David R.  

Mannheim, Karl  

Marris, Peter and Martin Rein  

Marx, Gary  

McAfee, Rosemary Ann  

McCall, George J. and J.L. Simons, eds.  

McGowan, Eleanor Farrar  

McKee, James B.  

Millman, Marcia and Rosabeth Moss Kanter, eds.  

Mitchell, E.C.  

Moore, Joan W.  

Mumford, Lewis  
Nader, Laura

Nagel, Stuart S., ed.

Nann, Richard C.

Nann, Richard C. and Penelope Stewart

Nash, Dennison and Ronald Wintrob

Neufeld, Edward Peter
1964 Money and Banking in Canada. Toronto: McClelland and Stewart.

New, Peter Kong-Ming and Richard M. Hessler

O'Connor, James

Oleson, Virginia L. and Elvi Waik Whittaker

Ontario Association of Housing Authorities
1964 Good Housing for Canadians. Toronto.

Ottaway, Richard N., ed.

Panitch, Leo, ed.
Parlett, Malcolm and David Hamilton  

Patton, Michael Q. et al.  

Patton, Michael Q.  

Pelto, P.J.  

Phillips, D.C.  

Porter, John  

Potter (Levitan), Alberta Z.  

Powell, Alan, ed.  

Quinney, Richard  

Rainwater, Lee  

Ranney, Austin, ed.  
1968 Political Science and Public Policy. Chicago: Markham.

Rein, Martin  

Rein, Martin and S.M. Miller  
Reynolds, Larry T. and Janice M., eds.

Riecken, Henry

Riecken, H.W. and R.F. Boruch, eds.

Rischin, Moses

Rist, Ray C.

Rivlin, Alice M.

Robbins, S.P.

Roethlisberger, F.S.

Rossi, Peter and Richard A. Berk

Rossi, Peter H., Howard E. Freeman and Sonia R. Wright

Rossi, Peter H. and Walter Williams, eds.

Rowbottom, Ralph
Saroff, Jerome R. and Alberta Potter Levitan

Schatzman, Leonard and Anselm Strauss

Schensul, Stephen L. and Jean J. Schensul

Schumacher, E.F.

Seeley, John R., L. Alexander Sim and Elizabeth Loosely

Shashkolsky, Leon

Shipnuck, Leslie, Dennis Keating and Mary Morgan

Shostak, Arthur B., ed.

Siegel, Karolynn and Pamela Doty

Sjoberg, Gideon ed.

Smith, Dorothy E.
Stake, Robert E., ed.

Stanford Evaluation Consortium

Stern, Rita and Jane Rosettis

Strathcona Property Owners and Tenants Association

Strauss, Anselm

Suchman, Edward A.

Thrasher, Jean H., Jack L. Franklin and Lee D. Kittredge

Treasury Board of Canada, Comptroller General

Treasury Board of Ontario

Trend, M.G.

Tresemer, David
Tuckman, Gaye

Vaughan, Ted R. and Larry T. Reynolds

Vidich, Arthur J., J. Bensman and M.R. Stein, eds.

Vivelo, Frank Robert

Wacaster, C. Thompson and William A. Firestone

Warren, Roland L.

Warwick, Donald P.

Warwick, Donald P. and Samuel Osherson, eds.

Waters, W.G.

Webb, Eugene J., Donald T. Campbell, Richard D. Schwartz and Lee Sechrest

Weber, George and George McCall, eds.
Weiss, Carol, ed.

Weiss, Miriam J.

Wheeler, Stanton

Wheelwright, Edward L. and Kenneth D. Buckley, eds.

Whittaker, Elvi

Wholey, J.S.

Wildavsky, Aaron

Wilson, James Q., ed.

Wiseman, Jacqueline P.

Zimmerman, Don H.
APPENDICES
APPENDIX 1: DESCRIPTION OF THE RESEARCH TEAM: BACKGROUND AND RESPONSIBILITIES

The research team was set up in such a way that the input from a number of different disciplines could be included in the final product. The staff of the group involved in the SRP Evaluation, Stage II, was composed of the following people:

1. Project Co-ordinator: Laurie Hurwitz, whose main background is in social and community planning both for governments and community groups in British Columbia and Ontario.

2. Research Director: Alberta Levitan, a sociologist who has worked in municipal planning and community organizing in addition to extensive work in the design and implementation of surveys.

3. Cultural Anthropologist: Gloria King, who has studied and worked in the Chinese communities of Johannesburg and London, is herself Chinese, and is familiar with the basic community and kinship groupings of the Strathcona area.

4. Field Co-ordinator: Jonathan Lau, a social worker who is originally from Hong Kong, and who has been involved in the Strathcona community for more than nine years, was on the staff of the SRP, is fully bilingual, and supervised all the Chinese interviews.

5. Land Economists: David Baxter and Dan Ulinder, both from the Faculty of Commerce, University of British Columbia, whose expertise in land development and housing markets was instrumental in developing a critical analysis of the economic ramifications of the SRP.

6. Physical Planner: Marlene Hier, a planner teaching at the Faculty of Engineering, University of British Columbia, who examined the upgrading of engineering services through the SRP.

7. Data Processor: Michael Goldberg, a social worker whose extensive background in processing and analysis of social science research data was advantageous in the coding and analysis of information from the interview schedules.

8. Technical Producer: Hayne Wai, formerly SPOTA executive director, experienced in research, media work and graphic design, who assisted in all technical details of producing this report.

9. Planning Intern: Suzanne Huzel, a student in the School of Community and Regional Planning, University of British Columbia, who assisted in the analysis of data from the interviews with social service agency workers.
We had additional assistance from a number of other part-time workers to do interviewing and coding. These included: Myra Eadie, Linda Chow, Mei Lai Wong, Susan Chang, and Penny Gurstein.

Source: Adapted from Alberta Levitan and Laurie Hurwitz, Strathcona Rehabilitation Project Stage II Evaluation (Vancouver 1977) p. 151.
APPENDIX 2: RESEARCH STRATEGIES: METHODS OF INQUIRY AND DATA SOURCES

A. Methods

This second stage of the evaluation of the Strathcona Rehabilitation Project was intended to provide a comprehensive analysis of the impact of the project from a number of differing points of view. In order to construct that type of analysis it was necessary to have data from all the various actors who played significant roles in the development and implementation of the project. In addition, we had to obtain information on the major components of the project: physical, social, economic and political.

Therefore, the research design involved a process of collecting and assimilating data from a variety of sources and of different types--on a scale from mainly qualitative to mainly quantitative. The following gives a brief listing and description of the types of data collected.

B. Data Sources

1. Selected information from the CMHC files of all (229) approved applicants for the grant/loan, including basic financial and procedural transactions between participants and the various members of the Site Office team, also the listing, details and costs of work done to rehabilitate each dwelling.

2. A series of 30 focused interviews (taped) with representatives of the three levels of government (CMHC, Provincial Housing Department, City Planning Department), and the community group which was the fourth partner to the agreement (SPOTA) as well as members of the Site Office staff.

3. A random sample survey of 49 participants in the housing rehabilitation program using a structured interview schedule developed in English and translated into Cantonese (using interpreters for those respondents who were Italian speaking). Predominantly owners of single family homes were interviewed.

4. A quota sample of 15 participants and/or non-participants who were mainly owners of multiple dwelling unit structures. They were interviewed using a somewhat modified form of the interview schedule developed for the random sample.

5. Field notes from conversations with a cross-section of residents, members of the SPOTA executive, provincial and federal civil servants who have played various roles in the area, local community service workers and others who have at some recent time lived or worked in the area.
6. Data from a survey of 105 social service workers in seven social agencies. The survey was concerned with community needs and problems including a set of questions on knowledge and evaluation of the SRP. The overall survey was designed by Michael Goldberg, a social worker hired by the Strathcona Community Resources Board, and the section on the impact of the SRP was designed by our study team. Interviews were done by students from a research class at the School of Social Work, University of British Columbia.

7. Field notes from conversations with staff of the engineering department of the City of Vancouver, as well as relevant financial material and information on engineering improvements in the area.

8. Data on all (441) real property sales transactions within the SRP boundaries from 1965 to the end of the third quarter of 1975. These were categorized as to whether or not the property in question was purchased by the City of Vancouver or by a private individual.

9. Major relevant documents were used in the analysis (see Bernard 1973; Bell and Moore 1975; Cahn and Passett 1974; Chan 1971; Cole 1976; Fraser 1972; Gans 1962; Goldberg 1976; Lowden 1972; Nann 1971; Spearing 1970; SPOTA 1973; Wai 1976).

Source: Adapted from Alberta Levitan and Laurie Hurwitz, Strathcona Rehabilitation Project Stage II Evaluation (Vancouver 1977) pp. 149-150.