

Supervisor: Dr. Christopher E. Hodgkinson

ABSTRACT

Turbulence in the British Columbia education system encouraged Ministry and School District administrators to develop a holistic, formative evaluation process for school districts. Called the Information Profile System (IPS), the process was intended to assist district administrators with improvement of accountability, professional teamwork and commitment, decision making, and gain of public support for education. However, the literature also suggested that while formative evaluation holds great theoretical promise for administration, the practical results are often disappointing. Examination of a first pilot confirmed this viewpoint. The central problem of the study therefore became to ascertain whether the IPS could be refined and reformatted to act as an effective leader-substitute in British Columbia; for the literature on both administration and evaluation suggested that the IPS could be considered a "leader-substitute" evaluation process, a series of tasks, procedures, and processes intended to enhance leadership effectiveness through stakeholder involvement in evaluation.

To solve the problem action research employing the IPS in evaluations of three school districts was conducted during the 1988-89 school year. Interviews of stakeholder participants in each district evaluation and other data were naturalistically evaluated in order to solve the main problem and three sub-problems: (1) how could the IPS be refined to improve the achievement of its goals; (2) what factors limited the IPS's effectiveness; and (3) what insights can be gathered into the leader-substitute construct of leadership? The IPS procedures and processes are described and critiqued in three district evaluation case studies. Meta-evaluation of the cases produced findings related to the restructuring of the IPS, its effectiveness, and the role of formative evaluation in administration.

The research suggests that a substantially reconstituted IPS can assist administrators with accountability, but only marginally effect the other purposes envisaged for the evaluation unless ownership and follow-through amongst the larger community of stakeholders is developed. The research also suggests that senior administrator commitment, moral fibre, and management of meaning skills are major factors limiting the success of formative evaluation. Finally, the research suggests that the "leader-substitute" construct of leadership has some conceptual merit as a characterization of the IPS, and for the dynamics of leadership; a characterization commensurate with a "subjectivist", or "humanistic" view of administration.

Examiners:

iii

~~Dr. C. E. Hodgkinson, Supervisor (Department of Communications & Social Foundations)~~

~~Dr. P. O. Evans, Departmental Member (Department of Communications & Social Foundations)~~

~~Dr. V. J. Storey, Departmental Member (Department of Communications & Social Foundations)~~

~~Dr. J. Jackson, Outside Member (Department of Physical Education)~~

~~Dr. D. Chabassol, Outside Member (Department of Psychology)~~

~~Dr. S. Mussio, Outside Member (Ministry of Education)~~

~~Dr. A. MacKay, External Examiner, (Department of Educational Administration)~~

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION.....	1
<i>Introduction</i>	1
The Information Profile System.....	2
Holistic Evaluation.....	3
The IPS: Holistic Evaluation as a Leader-Substitute.....	4
<i>Statement of the Problem</i>	4
The Pilot Study.....	6
Purpose of the Study.....	7
1. Refinement and critique of the IPS model.....	7
2. Understanding formative evaluation's limitations in school districts.....	8
3. Insights into administrative limitations of formative evaluation, and the leader-substitute construct.....	8
Definitions.....	9
Limitations.....	10
<i>The Significance of the Study</i>	11
1. Conceptual significance.....	11
2. Practical Significance.....	12
<i>Summary</i>	13
 CHAPTER II: LITERATURE REVIEW.....	 14
<i>Part I: Background</i>	14
<i>Education Politics in British Columbia</i>	14
The Administrative Crisis in Education.....	15
<i>The Information Profile System</i>	17
Stage 1.....	18
Stage 2.....	19
Stage 3.....	20
The IPS in 1988.....	20
The Need for Further Refinement.....	20
The Spring, 1988, Field Test.....	21
<i>Summary</i>	21

Part II: The Literature on Administration.....	22
<i>The Nature of Knowledge and Theory in Educational Administration.....</i>	<i>22</i>
Philosophical Premises of the Emerging Paradigm.....	23
The Impact of the Emerging Paradigm on Educational Administration.....	24
Applying the Emergent Paradigm to This Dissertation.....	28
Part III: Leadership/Administration Theory and Organization Theory.....	29
<i>The Leader-substitute Construct.....</i>	<i>29</i>
The Role of Power and Leader-substitutes.....	30
"Ideological" Power and "Disciplinary" Power.....	32
The Leader-Substitute Construct.....	33
The Leader-Substitute Construct and Administration.....	34
<i>Effective Administration in Loosely Coupled Organizations... 36</i>	<i>36</i>
Principles of Effective Administration.....	38
The Role of Values in Moral Administration.....	39
Developing a Values Focus.....	42
The Role of Metaphor in Administration.....	48
The Educational Organization as Machine.....	48
The Educational Organization as Brain.....	49
An Educational Organization as a Holographic System.....	52
1. Redundancy of functions.....	53
2. Requisite variety.....	54
3. Minimum critical specification.....	54
Part IV: The Evaluation Function of the Information Profile System.....	55
<i>What is Evaluation?.....</i>	<i>56</i>
The Difference Between Evaluation and Measurement.....	56
Evaluation Research as an Administrative Tool.....	57
Values and Evaluation.....	58
<i>The Epistemological Debate in Evaluation.....</i>	<i>60</i>
Independent versus dependent reality.....	60
<i>Qualitative Strategy and Naturalistic Inquiry.....</i>	<i>63</i>
Naturalistic Inquiry.....	64

The Naturalistic Approach to Evaluation.....	65
Trustworthiness in Naturalistic Studies.....	66
Stakeholder Participation in Naturalistic Evaluation.....	68
<i>Evaluation Models and Qualitative Strategies.....</i>	68
Evaluation Models.....	68
1. The CIPP model.....	71
2. Formative and summative evaluation.....	72
Evaluation Models Compatible with Qualitative Strategy..	73
1. The transaction model.....	72
2. The goal-free model.....	72
3. The decisionistic model.....	73
<i>Purposes for Educational Evaluation.....</i>	74
Conflict Reduction and Complacency Reduction.....	74
Evaluation for Accountability.....	75
Evaluation for Understanding.....	76
Evaluation for Gain of Public Support.....	77
<i>Utilization-focused Evaluation.....</i>	78
Patton's Utilization Guidelines.....	78
Factors Affecting Utilization.....	79
<i>Factors Limiting Evaluation Utility and Influence.....</i>	81
<i>Holistic Evaluation as a "Utilization-focused" Strategy.....</i>	84
What is Holistic Evaluation?.....	85
The Holistic Evaluation Process.....	85
1. Values, administration, and holistic evaluation.....	88
2. Qualitative and quantitative data collection.....	88
3. The holistic design versus evaluation research.....	89
4. Holistic design and stakeholder participation.....	90
5. Holistic design and key indicators of performance....	91
Factors Limiting the Use of Holistic Evaluation Methods....	92
Standards of Practice for Program Evaluations.....	93
<i>Summary.....</i>	96
CHAPTER III: DESIGN AND METHODOLOGY.....	07
<i>Action Research and Naturalistic Inquiry.....</i>	97
<i>The Problem.....</i>	98
1. Refinement and Critique of the IPS.....	100
2. To Understand the Limitations of Formative Evaluation in School Districts.....	100

3. Formative Evaluation, Administration, and the Leader-substitute Construct.....	101
<i>Strategy and Procedures</i>	101
1. Refining the IPS.....	101
2. Determining the Effectiveness of the IPS.....	103
(a) accountability.....	103
(b) Teamwork and commitment.....	103
(c) Decision making.....	105
(d) Public support.....	105
3. Formative Evaluation's Limitations and the Leader-substitute Construct.....	106
<i>Sample</i>	106
<i>Data Collection Techniques</i>	107
<i>The Case Study Approach</i>	110
1. Generalizing from multiple case studies.....	111
<i>Validity and Reliability</i>	112
Trustworthiness.....	112
<i>Limitations and Delimitations</i>	114
Limitations.....	114
Delimitations.....	116
 CHAPTER IV: CASE STUDY NUMBER ONE.....	 117
 <i>The Inlet School District Evaluation</i>	 117
<i>The Context and Background to the Evaluation</i>	118
The Educational Context.....	119
<i>Description and Critique of the Evaluation Procedures and Processes</i>	123
1. Groundwork.....	124
2. The Stakeholder Collaborative Approach.....	125
3. The Four-Goal, Six-Attribute Framework.....	129
4. Key Indicator Development.....	130
5. Questionnaire/Survey Process.....	132
6. Data Collection, Analysis, and Interpretation Procedures.....	134
7. The Evaluation Report.....	135
<i>Applications of the Inlet District Evaluation</i>	136
Description.....	136
Critique.....	139
<i>Summary</i>	142

Chapter V: CASE STUDY NUMBER TWO.....	144
<i>The South Coast School District Evaluation</i>	144
<i>Context and Background to the Evaluation</i>	144
<i>Description and Critique of the Evaluation Procedure and Processes</i>	149
1. Groundwork.....	149
2. The Stakeholder Collaborative Approach.....	151
3. The Four-Goal Six-Attribute Framework.....	154
4. Key Indicator Development.....	155
5. Questionnaire/Survey Information.....	160
6. Data Collection, Analysis, and Interpretation Procedures.....	162
7. The Evaluation Reports.....	166
<i>Applications of the Evaluation Information</i>	168
a. Description.....	168
b. Critique.....	169
<i>Summary</i>	170
Chapter VI: CASE STUDY NUMBER THREE.....	171
<i>The Interior School District Evaluation</i>	171
<i>The Context and Background to the Evaluation</i>	172
The Educational Context.....	173
<i>Description and Critique of the Procedures and Processes of the Evaluation</i>	175
1. Groundwork.....	175
2. The Stakeholder Collaborative Approach.....	179
3. The Four-Goal, Six Attribute Framework.....	182
4. Key Indicator Development.....	184
5. Questionnaire/Survey Information.....	186
6. Data Collection, Analysis, and Interpretation Processes.....	188
7. The Evaluation Reports.....	190
<i>Applications of the Interior District Evaluation</i>	194
a. Description.....	194
b. Critique.....	196
<i>Summary</i>	199

CHAPTER VII: META-EVALUATION OF THE INFORMATION PROFILE SYSTEM.....	200
<i>Meta-Evaluation of the IPS Methodology</i>	200
1. The Four-Goal, Six-Attribute Framework.....	200
2. Key Indicator Measurements of the Goals and Attributes.....	203
3. Questionnaire/Survey Procedures and Evaluation Reports.....	208
4. Focusing Stakeholders on the Task of Evaluation.....	211
5. The Stakeholder Collaborative Approach.....	212
<i>Meta-Evaluation of the the IPS</i>	214
1. Accountability.....	214
2. Teamwork and Commitment.....	216
3. Decision Making.....	221
4. Public Support.....	224
<i>Summary</i>	225
 CHAPTER VIII: GENERAL DISCUSSION AND COMMENTARY ON EVALUATION AND LEADER-SUBSTITUTE THEORY.....	 226
<i>Evaluation</i>	226
1. Observed Limitations.....	226
a. Traditional limitations.....	226
b. Other limitations.....	228
2. Evaluation and Administration.....	232
<i>Leader-Substitute Theory</i>	245
Administrative Effectiveness.....	248
<i>Further Implications</i>	255
<i>Summary</i>	258
 CHAPTER IX: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS.....	 259
<i>Background</i>	259
1. Purposes.....	259
2. Procedures.....	259
3. Principles.....	260
<i>Study Purposes</i>	261

<i>Design and Methodology</i>	261
<i>Findings</i>	263
General Limitations.....	265
Evaluative Limitations.....	265
<i>Leader-Substitute Theory</i>	266
<i>Conclusions and Implications for Further Research</i>	267
<i>Questions for Further Research</i>	269
Further Recommendations.....	270
<i>Conclusion</i>	272
 REFERENCES.....	 273
 APPENDICES.....	 284
 <i>Appendix A1: The Four-Goal and Six Attribute Framework</i>	 284
<i>Appendix A2: Sample pages of an Information Profile</i>	288
<i>Appendix A3: Sample of one of the seven Questionnaires used in the Inlet and South Coast District Evaluations</i>	290
 <i>Appendix B1: Interior District Interview Protocol</i>	 292
<i>Appendix B2: Audit trail guide</i>	307
 <i>Appendix C1: Letter of agreement between Inlet School District and the P.E.R.B., Ministry of Education</i>	 310
<i>Appendix C2: Chronology of Events in the Inlet Evaluation</i>	312
<i>Appendix C3: Survey Administration Procedures in the Inlet School District</i>	314
<i>Appendix C4: Data Interpretation and Analysis Procedures in the Inlet School District Evaluation</i>	315
 <i>Appendix D1: Chronology of Events in the South Coast Evaluation</i>	 315
<i>Appendix D2: Survey Administration Procedures in the South Coast School District</i>	318
<i>Appendix D3: The Q-Sort process utilized to prioritize key indicators in the South Coast Evaluation</i>	319
<i>Appendix D4: Letter outlining Stakeholder Team responsibilities for the survey administration in the South Coast School District</i>	321
<i>Appendix D5: Data Interpretation and Analysis Procedures in South Coast School District Evaluation</i>	323

<i>Appendix E1: Chronology of Events in the Interior Evaluation.....</i>	324
<i>Appendix E2: Interior School District Mission Statement... ..</i>	326
<i>Appendix E4: Sample of Indicator lists form the Interior School District Evaluation.....</i>	328
<i>Appendix E5: Interior School District Survey Methodology and Design.....</i>	329
<i>Appendix F1: Frequency responses of stakeholder perceptions as to the positive value of the Four-Goal Six Attribute Framework, and its accuracy.....</i>	330
<i>Appendix F2: Frequency responses of stakeholders to the Development and use of district and ministry key indicators of performance: positive value and accuracy... ..</i>	331
<i>Appendix F3: Frequency responses of stakeholder perceptions with respect to questionnaire desirability and accuracy.....</i>	333
<i>Appendix F4: Frequency responses of stakeholder perceptions of evaluation descriptors re: efficiency and effectiveness of the evaluation processes and procedures.....</i>	334
<i>Appendix F5: Frequency of responses of stakeholders perceptions of descriptors re: information sharing qualities of the evaluations processes and procedures.....</i>	335
<i>Appendix F6: Frequency distribution of words chosen as three words that best describe stakeholders' role.....</i>	336
<i>Appendix F7: Frequency distribution of word categorization- describes role well (w), describes role somewhat (s), describes role not at all (n).....</i>	337
<i>Appendix F8: Frequency distribution of stakeholder perceptions of evaluation descriptors re: ethical quality.....</i>	338
<i>Appendix F9: Frequency distribution of stakeholder perceptions of evaluation descriptors re: authority relationships.....</i>	339
<i>Appendix F10: Frequency responses of Stakeholder Team members to possible achievements of the evaluation process.....</i>	340
<i>Appendix F11: Frequency responses of Stakeholder Team members to possible achievements of the evaluation process.....</i>	341

<i>Appendix F12: Frequency responses of Stakeholder Team perceptions of impact of the evaluation in terms of gaining public support for education.....</i>	<i>342</i>
<i>Appendix F13: Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.....</i>	<i>343</i>
<i>Appendix F14: The B.C. Government's Mandate Statement for Education.....</i>	<i>358</i>
<i>Appendix G1: Steps of Value judgement experienced in the first stage of the Inlet evaluation.....</i>	<i>360</i>
<i>Appendix G2: Employing Hodgkinson's Value Paradigm to characterize value priorities in the formative evaluation model.....</i>	<i>361</i>
<i>Appendix G3: Q-Sort ratings of stakeholder opinions of the main factors accounting for the success of the district evaluation order of priority.....</i>	<i>362</i>

Tables

Table 2.1: Greenfield's Analysis of social reality.....	27
Table 2.2: Procedures for ensuring trustworthiness.....	67
Table 2.3: Evaluation models.....	70
Table 2.4: An outline of the Utilization-Focused Approach to Evaluation.....	79
Table 2.5: Comparison of "Typical" SBE and SCE approaches.....	90
Table 2.6: Positive and negative features of the SCE approach.....	91
Table 2.7: Standards of good practice for Program Evaluation.....	94
Table 3.1: The Naturalistic Design Applied to the Meta-evaluation Study.....	99
Table 3.2: Demographic characteristics of Inlet, South Coast, and Interior School Districts.....	108
Table 7.1: The 12 words most frequently chosen as describing individual stakeholder roles in the three district evaluations: by district and overall.....	217

Figures

Figure 2:1: The two components of administration.....	36
Figure 2:2: The Value Paradigm.....	41
Figure 2:3: Developing an empirical vision.....	43
Figure 2:4: The “Double-loop” learning capacity of the brain.....	50
Figure 2:5: Artificially constructed learning loops.....	52
Figure 2:6: Principles of holographic design.....	53
Figure 2:7: “Value-free” reality versus “value-laden” perceptions..	60
Figure 2:8: The epistemological base of evaluation models.....	69
Figure 2:9: The three stages of holistic evaluation.....	86
Figure 8:1: Formative evaluation and learning loop for organizational renewal.....	233
Figure 8.2: Formative evaluation as the force moving the organizational learning cycle forward in an active- adaptive-reactive way.....	237
Figure 8.3: The components of administration: person-based leadership plus leader-substitutes.....	250
Figure 8.4: The shifting locus of administrative action: constantly in operation.....	250
Figure 8:5: The dimensions of administrative effectiveness.....	252

CHAPTER I

Introduction

A major challenge faces educational leaders today. Across North America demands are being made upon custodians of education to maximize the performance of schooling amidst an environment of disenchantment and declining resources (Doyle & Levine, 1986; in't Veld, Spee, Tseng, & Sandbergen, 1987; Murnane, 1987). In British Columbia, this is a particularly potent concern (Royal Commission on Education, Summary Findings, 1988). Demands for a higher quality education system are being made during a time of extended conflict in the educational arena (Coleman, 1985; B.C. Ministry of Education [BCME], 1987; Callam & Fleming, 1988). Difficulty in meeting these demands make school district superintendents and trustees, who are directly accountable to their public, particularly vulnerable (Storey, 1987; Fleming, 1988).

Concomitantly, the Ministry of Education needs to meet this challenge. It must reaffirm its responsibility to justify the efficacy and quality of provincial educational programs (McIver, 1985; Royal Commission on Education, Summary Findings, 1988). Yet how can these demands be met when no measures of "school performance" are clearly accepted by the educational community (Murnane, 1987)? In addition, administrators, charged with the responsibility for evaluating and maintaining quality, are forced to adjust to and understand new roles being thrust upon them (Bill 20, 1987; Bill 67, 1989). In severe cases, some are struggling to ensure their professional survival (Fleming, 1988; Storey, 1987). How can they, in the absence of quality performance measures, successfully communicate their successes and needs for assistance to the public?

Many writers have suggested that formative evaluation can assist educators in meeting this challenge (Stufflebeam, 1983; Scriven, 1983; Glasman, 1986). "Evaluation is a well accepted element in the management of education"

(MacPherson, Crowson & Pitner, 1986, p. 84). Formative evaluation can lead to improved decision making by governmental institutions, and also serve "functions of complacency reduction and conflict management; it may serve other functions well" (Floden and Weiner, 1983, p. 186). Glasman (1986) suggests that other functions are improvement, accountability, motivation and gain of public support, and the exercise of authority.

The Information Profile System

Consistent with this reasoning the Ministry of Education, in collaboration with the Association of British Columbia School Superintendents (ABCSS), decided to design a comprehensive evaluation system to assist them in dealing with the issues of improvement, accountability and public disenchantment with education (McIver, 1985). The Information Profile System (IPS) is the outcome of this collaboration (Program Evaluation Branch, B.C. Ministry of Education [PEB BCME], 1988b).

The Information Profile System is designed to assist with the leadership functions of direction setting, or "purposing" (Barnard, 1938; Hodgkinson, 1978, 1983; McIver, 1985; Storey, 1988), accountability (McIver, 1985; Glasman, 1986; McPherson, Crowson, & Pitner, 1986), and motivation and gain of public support (Glasman, 1986; Sergiovanni, Burlingame, Coombs, & Thurston, 1986). It does so by using evaluation, "the collection and use of information to make decisions about an educational program" (Cronbach, 1983, p. 101).

The IPS consists of four components (PEB BCME, 1988a). The first is a values framework that defines the system's purposes, both in terms of educational outcomes for students and instrumental goals for organizations. The second is a communications system that encourages the sharing of information on schooling between the Ministry, the school district, and the public. A third component is a series of "key indicators"--meaningful, appropriate, and practical measures--that can be used to "operationalize" each value. A fourth component is a "reporting out format". Samples of reports have been provided so administrators can effectively report out to fellow professionals, the school board, and the public. A final component is the provision of a guide book to

assist with interpretation of data, understanding of data, and planning for the future.

Holistic Evaluation

The overall evaluation process is an eclectic one, as a number of different evaluation designs are employed. Such a method has been termed "holistic" evaluation, since it "is open to gathering data on any number of aspects of the setting under study in order to put together a complete picture of the social dynamic of a particular situation or program" (Patton, 1980, p. 40). Moran (1987) describes its purpose as follows:

...This holistic model is designed to provide managers with information on how policy choices can affect the efficiency or effectiveness of their programs....The utility of this approach lies in the comprehensiveness of the research design, which provides policymakers with sets of data on how well a program is doing from bureaucratic, political, and societal perspectives (p. 613).

Although holistic evaluation employs a mix of different methodologies and designs, the naturalistic methodology (Lincoln & Guba, 1985) employed wholly or partially appears particularly suited to the evaluation of complex social entities such as education. This is because evaluation of large education systems cannot be done except within a very value-laden context (Holmes, 1987). If the purpose of an evaluation is understanding, in a "hermeneutic" sense (Smith, 1983), or "a decision to do something" (Nilsson and Hogben, 1983, p. 92) then a "value-based" versus a "value-free", qualitative approach to evaluation is required (Scriven, 1983a). The naturalistic methodology is predicated on this belief.

The naturalistic methodology can use both qualitative and quantitative methods (Lincoln and Guba, 1985), but stresses multiperspectives (Farley, 1987), participation of stakeholders (Ayers, 1987; Greene, 1987a), and value-pluralism (Greene, 1987b). As holistic evaluation must concern itself with the values and aspirations of members of the educational community, the use of such methods are consistent with its intent.

The IPS: Holistic Evaluation as a Leader-Substitute

The purpose of the IPS is to assist formal leaders with administration of education during a time when they are in need of justifying and improving the quality of education (Auditor General's Report, 1987). It is intended to be a form of substitute leadership--leadership embedded in a formative evaluation structure and process--to "act in the place of" a specific leader behavior" (Howell and Dorfman, 1981, p. 715).

An underlying assumption of both the IPS and of the leader-substitute theory is that power can be ensconced in organizational characteristics as much as in individuals (Kerr & Jermier, 1978). In traditional leadership theory, effective leaders are able to influence subordinates by dint of leadership traits, or personality (Burns, 1978; Halpin & Winer, 1957; MacGregor, 1960). Situational leadership theory makes the same claims, suggesting only that leader traits are somewhat mitigated by the context within which they operate (McPherson, Crowson and Pitner, 1986). The leader-substitute theory states, however, that the leadership function need not reside in the person designated to lead. Recent studies (Freeston, 1987; Gamoran and Dreeben, 1986; Kerr and Jermier, 1978; Pitner, 1986) indicate that organizational characteristics relating to task, subordinates, and structure can act as "neutralizers" of formal leadership behavior.

Processes and procedures within an organization can assume a leadership role by directing subordinates--to the degree that they can enhance or replace "the leader's ability to influence subordinate satisfaction and performance" (Freeston, p. 46). A logical extension of this proposition is that structures, tasks, and subordinate characteristics utilized by the Information Profile System, and designed according to effective leadership principles, can enhance an administrator's ability to meet educational challenges.

Statement of the Problem

Earlier it was stated that holistic evaluation, employing some forms of naturalistic methodology appears, theoretically, to be of value in dealing with issues of accountability, improvement, and controversy in the educational

arena. As a result educational leaders in British Columbia have designed the IPS to achieve these goals. Yet, despite its theoretical promise, holistic evaluation has not--in practice--been as successful as it might be. Indeed, "the use of evaluation as a technique to control educational organizations has not been numbered among the strongest of administrative tools" (McPherson, Crowson & Pitner, p. 84). Alkin, Daillak & White (1979), agree, commenting that there is a "great dissatisfaction with evaluation's lack of usefulness and impact" (p. 14). A pilot study of the IPS in twelve school districts during the 1986-87 school year verified this observation with respect to British Columbia (Mussio, 1987).

Clearly, then, the problem is to design a holistic evaluation process that works, and that provides provincial and school district administrators with a methodology that enhances their leadership capacity. This dissertation approaches the problem from two points of view. The first is to examine, in detail, the literature relating both to effective administration and effective evaluation, and, in light of the leader-substitute construct, design a theoretical model that will suggest tasks, structures, and subordinate characteristics that will enhance the potential for holistic evaluation to become a success within the school district. The second is to employ this model, in a form of action research, throughout a number of districts in the province, and refine it in light of actual practice.

The IPS is the starting point in these efforts. It is currently in its third year of development; it is a leader-substitute whose design is still emerging. Some might expect that after three years of conceptual growth, field testing, and subsequent refinement the IPS would be well established in the field. This is not the case (Mussio, 1987). A number of factors are responsible for this situation, not the least of which is the simple fact that the breadth and purpose of the project--when one considers the political climate in B.C., the desire to produce a model that is applicable to 75 school districts and 28,000 teachers, and the non-reliability of outside models (Stufflebeam & Welch, 1986)--is huge. In addition, some of the natural impediments to evaluation (Scriven, 1983a; Patton, 1980, 1983; Stufflebeam & Welch, 1986) must be overcome. In essence, then, the problem is to revise an existing evaluation model so as to enhance its

utility, integrity, and effectiveness in achieving the goals of accountability and district improvement.

The Pilot Study

The first practical step in the IPS's development was a pilot study held in the 1986-87 school year in twelve school districts across the province. Evaluation of that study revealed some limitations to its effectiveness (Mussio, 1987). These weaknesses reflect both conceptual and practical problems with its methodology, and concerns regarding the process of implementation.

Methodological concerns related to the gathering and disseminating of information. A process that would distinguish valuable information from insignificant information, and then package it so as to be understood by a variety of different audiences with different knowledges and perspectives, needed to be designed. In addition, few administrators or other leaders have any formal training in evaluation procedures beyond the classroom level, nor with respect to using evaluation information for action planning for the future. Detailed explanations or guidance in these areas was required.

Implementation needs related to overcoming the natural resistance and inertia of many different constituent groups in the process. Evaluation is naturally resisted, (Scriven, 1983) and particularly so when it is perceived as being used by formal leaders as a political tool (Stufflebeam & Welch, 1986), a problem that was exacerbated by the political climate surrounding education in B.C. In addition, practitioners do not necessarily see the value of evaluation, nor do they have the expertise to carry it through (Moran, 1987). Resources (time and money) have not been traditionally allocated to this kind of project. Thus, the Information Profile System is much more difficult to design than a simple evaluation system, and much more problematic to implement than most leadership initiatives.

Yet the same pilot project provided insights as to the theoretical potential of evaluation at the school district level--even though this potential was unrealized, or only partially realized in practice. It was for this reason that a decision was made to redesign the IPS so as to overcome the factors limiting its effective utilization, and in so doing, gain a better understanding as to its true

effectiveness in dealing with the problems of improvement, accountability, and gaining of public support for education. The problems of methodology refinement and implementation had to be dealt with in tandem, not isolation; thus the two-fold approach of melding conceptual design with action research was suggested as a mechanism to enhance potential success.

To this end the IPS has been refined and reformed (PEB, BCME, 1988b). A second field test will be conducted in the spring, fall and winter of 1988-89, in seven school districts across the province. Three in-depth case studies will form the substance of this dissertation.

Purpose of the Study

The results of the pilot project created three purposes for this study:

1. Refinement and critique of the IPS model. By involving the researcher in the three individual evaluation processes, and by adjusting practice while "in action," an effort was made to discover better methodologies and implementation strategies for the IPS that enhanced its utility and effectiveness within each school district. Also, a meta-evaluation of the three district evaluations, using the recorded empirical observations and experiences of the action research, was conducted to refine the structure and process of the IPS model.

A second stage in the achievement of this purpose was a meta-evaluation, in a summative sense, of the overall effectiveness of the IPS model as a administrative tool for educational leaders. Effectiveness was measured in the domains of accountability, teamwork and commitment, decision making, and public support for education. Estimates of these were qualitatively determined from information gleaned from interviews and records, specifically relating to: (1) an increase in short and long range plans, and the quality of such plans; (2) an increased sense of motivation, teamwork and commonality of purpose amongst participants, revealed through a greater willingness to become involved in implementation of these plans; and (3) better information flow and understanding of the education system by members of the education community.

2. Understanding formative evaluation's limitations in school districts. The same meta-evaluation was used as an empirical data source from which to gain insights into the factors that limited the use of formative evaluation as a tool for school district improvement. In this section the observed limitations in the three case studies will be interpreted in light of the literature on evaluation limitations to be conducted in Chapter V. These include the political nature of evaluation and evaluation utility (Stufflebeam & Welch, 1986; Patton, 1980; Eisner, 1978); and practical limitations to evaluation effectiveness (Alkin, Daillak and White, 1979).

3. Insights into administrative limitations of formative evaluation, and the leader-substitute construct. A final purpose of the study is to gain further insight into the contention that formative evaluation seems to hold great promise as a tool for system improvement, but rarely is as effective as one would like it to be. As a naturalistic study, the intent is to realize the two purposes presented above but also to improve understanding of the link between evaluation and administration, the leader-substitute construct itself, and the factors that limit the usefulness of evaluation as an administrative tool.

Some of the concepts that need to be examined in order to achieve this purpose are the role of metaphor and vision in evaluation, as applied to the educational organization (House, 1983; Morgan, 1987); evaluation's relationship to issues of will and ethics (Barnard, 1938); philosophy, values and decision making (Hodgkinson, 1983); management of meaning and communication in educational organizations (Sergiovanni, Burlingame, Coombs & Thurston, 1987); the loosely coupled nature of organizations (Weick, 1985); problem-solving (Liethwood, 1983; Morgan, 1987); power (Hiley, 1987); vision (Storey, 1988); and the leader-substitute construct itself (Kerr & Jermier, 1978).

A more detailed discussion of the design, methodology, and instrumentation of the study will be presented in Chapter III. Before attempting to do so, however, the terms used in the paper need to be defined, and the limitations and significance of the study laid out.

Definitions:

1. **Leadership:** the exercise of the capacity to lead; to cause followers, through direction or influence, to pursue a goal.

2. **Administration:** the exercise of leadership within an organizational context. Administration is a sum of the capacity of both the "status" leader and leader-substitutes within an organization.

3. **Leader:** the person from whence the capacity to lead emanates.

4. **Administrator:** the individual in an organization occupying a "status" leadership role.

5. **Leader-substitutes:** The characteristics of subordinates, of task, or of structure within an organization that exercise the capacity to lead by utilizing person based, ideological or discipline based power.

6. **Holistic Evaluation:** An eclectic process of evaluation that utilizes a variety of research designs and methodologies to create a better understanding of the social dynamics and procedures taking place in large, complex social systems such as education.

7. **Naturalistic Evaluation:** An evaluation methodology that is predicated on the value-laden nature of reality, and that is devoted to understanding (hermeneutics), inductive development of theory, in-depth analysis of the dynamic relationship between variables, and the paramountcy of intuitive knowledge as opposed to empirical objectivity in research design.

8. **Meta-evaluation:** The evaluation of evaluation: the summative and formative judgments across case studies as to the overall success of an evaluation.

9. **The Information Profile System (IPS):** A holistic evaluation model designed to assist educational leaders in B.C. to deal with issues of district decision making, accountability, and gain of public and professional support.

10. The Four-Goal/Six-Attribute Framework: The Ministry of Education's mandate statement for the B.C. education system, whereby schools are expected to provide learning experiences to enhance student intellectual, social, human and vocational development; and school organizations are to exhibit the attributes of accessibility, relevance, cost effectiveness, quality management and accountability, professionalism, and public satisfaction.

11. Key Indicators: Measures of educational performance that are both qualitative and quantitative in design: qualitative, in that they are reflections of educator opinions as to how educational success can be measured; quantitative in that they are statistical statements of some desirable educational result.

Limitations

1. A purposeful sampling technique does not allow for transferability of the findings to any other context unless the reader deems the circumstances of the case study so described to allow transfer (Lincoln and Guba, 1985).

2. The study operates on a constrained timeline. A longitudinal study would be better than a year long research; the outcomes of success might not manifest themselves in this short time frame.

3. There was difficulty at times to get full cooperation from district representatives due to limitations on the ability to provide full confidentiality to participants; in addition, as Skrtic (1985) warns, the degree of trust established in each district varied. Nor was there unlimited access to people for reference checks.

4. The possible existence of a "Hawthorne Effect" (Hanson, 1985) that creates two possible limitations (this is discussed further in Chapter III). The first is related to the fact that the researcher assisted with the district evaluations. The second is due to the novelty of the IPS model itself. Adherence to the rigors of good naturalistic inquiry control the first; and the second is managed through awareness and factoring it out during the analysis stage of the evaluation.

The Significance of the Study

There are two categories of potential significance for the study: conceptual significance, and practical significance.

1. Conceptual significance

The concept of leader-substitutes has only recently been examined and discussed in the literature, beginning with Kerr and Jermier's study in 1978. Inherent in this concept is the belief that leadership--the ability to purpose, to communicate, and to empower subordinates--need not reside in the hands of the formal leader of the organization. Leader-substitute theory states that these qualities can also reside in an organizational structure, to the degree that they actually "neutralize" formal leader behavior (Freeston, 1987).

This study contends that the IPS, structured as a formative evaluation, can serve as a leader-substitute for the educational organization. If the IPS procedures and processes can be designed to create influences so that the personal evaluative skills of the administrator are enhanced by structures, subordinate characteristics and tasks inherent in the model, then an argument for the validity of the leader-substitute construct exists. This may lead to a clearer understanding as to the dynamics of effective administration, and in particular to the concept of situational leadership theory. If, in fact, the IPS model accomplishes the purposes of better accountability, gain of public support, and decision making for improvement, then an argument supporting the leader-substitute concept can be mounted.

This is significant because it implies that a way to reduce fluctuations in administrative ability is through the use of well-designed structures to address areas of weakness. The "use of evaluation as a technique to control educational organizations has not been numbered among the strongest of administrative tools" claim McPherson, Crowson & Pitner, (1986, p. 84). If the IPS can demonstrate a process to improve this circumstance, (albeit one idiosyncratic to the British Columbia situation), then others may see a similar potential in other venues. Other conceptual knowledges and practical knowledges may also be assisted by this study.

A second source of conceptual value relates to understanding the role of evaluation in administration. Stufflebeam & Welch (1986) state that "meta-evaluations are of utmost importance in helping to provide direction for improving evaluation services" (p. 166). As idiosyncratic as the case studies are to the districts they are to serve, the meta-evaluation component of the study may assist in the development of at least a better understanding of evaluation as an administrative tool.

2. Practical Significance

The practical value of this study on the IPS relates to the general topics of leadership and evaluation. Traditional leadership--as exercised by individuals in positions of authority in education in British Columbia--has been severely pressed by the public to justify the quality of schooling. If the IPS model can assist administrators by improving communication, accountability, and assisting in the development of long range plans for district improvement, it will help restore a sense of purpose and control to the education system. The roles of accountability, gain of public support, and decision making for improvement, if effectively provided for in the design of the system, will serve to augment the personal skills of administrators in running their organization. As Weick (1985) stated, reducing ambiguity helps reestablish control over the organization and contributes to a climate for innovation. An effective IPS model could be of much practical value to administrators if it can be applied to their individual circumstance.

A second domain of practical significance relates to the major challenge facing educational evaluators with respect to measurement of educational performance. Goals such as "human development", and "social development", so dear to the educator, have traditionally been immeasurable. As a result, even the most recent publications on educational effectiveness measure educational success in terms of student academic achievement and cost effectiveness (U.S. Department of Education, 1987) and tend to ignore the other major purposes of schooling. If this study can show that the goals of social and human development can be measured in a fashion acceptable to the educational community of the participating districts, then at least the

potential for this strategy has been established. The development of key indicator measures for human and social development has important implications for the management of education (Hanushek, 1987; Murnane, 1987).

Summary

This dissertation contends that the IPS, structured as a process of formative evaluation, can serve as a leader-substitute for the educational organization. It does so by providing a series of steps and processes, embedded in task, structure and subordinate characteristics that apply evaluation principles to administrative action. In so doing school district leaders can improve knowledge about education in the school district, revitalize constituent commitment, and provide direction for future improvement.

CHAPTER II

LITERATURE REVIEW

Part I: Background

Education Politics in British Columbia and the Information Profile System

Circumstances in the British Columbia education system today endow the issue of an "leader-substitute evaluation system" with particular significance. Educators find themselves enmeshed in an educational process that has become extremely politicized and turbulent (Calam & Fleming, 1988). The resulting ambiguity in goals and purposes for education challenge administrators by creating a need for new solutions to the problems of accountability, educational improvement, and motivation of public support. The Royal Commission on Education in British Columbia (1988) expressed this need in the following recommendations:

18. That the Ministry of Education, school districts, and schools take steps collaboratively to improve their accountability processes. Further, that in doing so, the following suggestions made to the Commission be considered:

- (1) the development of outcome measures related to established educational objectives;
- (2) the establishment of methods for relating outcome data to financial data and use of resources; and
- (3) the initiation of improved and more extensive means of periodically communicating such data to constituents.

19. That the Ministry of Education support district accountability initiatives with increased financial resources and expertise (p. 184, 1988).

The IPS is designed to assist in these regards.

The Administrative Crisis in Education

Two historical factors have contributed to a crisis in administration in British Columbia's education system. The first relates to the controversy engendered by the government's policy of financial cutback, or restraint, during the 1982-87 period, and the subsequent restructuring of funding for education (Calam & Fleming, 1988). A second, and contingent reason, relates to the negative relationships that exist between the British Columbia Teachers Federation [BCTF] on one hand, and the Government of British Columbia on the other. Relations between the two have been characterized by distrust and disrespect due to continual and fundamental disagreement about major educational and economic issues such as school roles and education funding (Legislature Staff, 1987; see Note 1). An outward manifestation of this unrest was the work-to-rule campaign conducted province-wide at the beginning of the 1987-88 school year (Parsons, 1987).

Another factor is the ongoing debate in the media concerning the quality of educational programs. While this dialectic has not necessarily been destructive, it has rarely been conducted in a rational, informed, or knowledgeable manner. Indeed, it has often been characterized by emotion, bitterness, and self-interest. As Calam and Fleming (1988) state:

...preoccupation with political conflict has exerted a paralyzing grip on provincial educational affairs, pitting labour and management, government and teachers, administrators and their staffs, one against the other to the detriment of all and especially to the disadvantage of school children (p. 52).

This has resulted in the success of the system being called into question, by both politicians and public (Bula, 1987c). They want to know how well students are achieving; they want to know how effective special education programs are; they want to know if things are getting better or worse; they want to

Note 1: Whereas this distrust is not necessarily directed at the Ministry of Education, as a separate entity from government, much ennui and skepticism as to Ministry intent does interfere with the development of a true collaborative relationship between the Ministry and the Teachers' union.

know how B.C. measures up with other educational systems. The response of educational leaders to these concerns has been "fragmented, inconsistent and often defensive" (Mussio, 1985, p.1).

The resulting public debate on education has affected superintendents, trustees, and principals in two main ways. The longevity of tenure of senior administrators has been drastically reduced (Fleming, 1988; Storey, 1987), and continuity in administration has suffered. As a result the ABCSS called for "a climate of stability and co-operation to permit responsible long range planning" (Calam & Fleming, 1988, p. 51).

In addition, all groups have found their ability to affect change severely hampered by the conflict between the BCTF and the government (Bula, 1987a). Administrators expend a great deal of energy not only learning new responsibilities, but also dealing with unproductive conflict far removed from the classroom (Bula, 1987b).

At the same time the Ministry of Education has been chastised for exercising poor leadership of the system. An external consulting team described its management practices as disjointed, lacking in trust, and suffering from poor communications with practitioners in the field (McIver, 1985). The same team also stated that the Ministry was bereft of complete planning and evaluation functions, and "lacking in leadership and definition of a total framework" (McIver, 1985, p. 1).

To alleviate these concerns, superintendents and Ministry officials decided in 1985 to work together to design an information-evaluation framework that would assist educators in answering the question, "how well are schools performing?" (Mussio, 1985, p. 3). Three goals were established:

1. To define school system expectations and outcomes;
2. To heighten public awareness of the achievements of education, in terms of its cost effectiveness, relevance, professionalism, and quality management; and
3. To measure school system accomplishments according to key indicators of outcome performance (McIver, 1985).

This administrative and evaluation initiative is now known as the Information Profile System. It was an attempt to reverse the administration crisis, and restore public stability to a "highly volatile and confrontational environment" (McIver, 1985, p. 3).

Early in 1987 a third factor--recent legislation (Bill 19 and Bill 20, 1987)--restructured roles, responsibilities, and relationships between teacher and principal, teacher and school board, and principal and school board. An unanticipated result has been the revitalization and unionization of the British Columbia Teacher's Federation, armed, for the first time in its history, with the legal right to strike (Smith, B., 1988). This has further contributed to administrative uncertainty. Indeed, protracted negotiations between school boards and teacher unions have been the overwhelming issue facing superintendents during the first half of the 1988-89 school year (N. Thiessen, personal communication, December 5, 1988). Educational administrators have only so much energy to expend, and such controversy takes time and energy away from improvement of the schools. Thus the rationale for the existence of the IPS is even more urgent today. Indeed, the recent Royal Commission in education states as one of its prime objectives the need to enhance productive collaboration between stakeholders in the educational arena (Royal Commission Summary Findings, 1988).

The Information Profile System

The development of the Information Profile System has proceeded through three stages.

Stage 1

The initial stage was its theoretical conception. The purpose of the project was described as follows:

1. To develop policies defining roles and responsibilities of schools, districts, and the Ministry in the evaluation of school programs.

2. To identify key indicators that can be used by various levels of the system to help determine to what extent goals of education (intellectual, social, human, vocational) are being achieved.

3. To select both "outcome" indicators (i.e., test scores, student attitudes, postsecondary enrolment patterns) as well as "contextual" or "process" indicators (retention rates, public attitudes, financial data, etc.).

4. To identify what data should be collected and reported at the school, district, and provincial levels (Mussio, 1985).

The originator, Dr. Jerry Mussio, summed up the purpose of the project as follows:

It is our intent that the indicators will be used in the process of communicating with our elected representatives and to the public on the state of education. We also believe that the evaluation system of the type being proposed here should be an integral component of planning and decision making at all levels and should form the basis of goal based discussions. Finally, it is our hope that our proposals can be used to help restructure provincial and local information systems (Mussio, 1985, p. 2).

During this stage the Four-goal/Six-attribute framework was established to describe a quality education system. Based on the Let's Talk About Schools (1985) exercise and an indepth analysis of recent literature on effective schools, the framework encapsulated the basic values of education for educational organizations in British Columbia. The four goals are: intellectual, social, human and vocational development. The six attributes of quality education systems are: accessibility, relevance, management and accountability, cost effectiveness, professionalism, and meeting public expectations. These are described in greater detail in Appendix A.1.

Stage 2.

In the summer of 1986, twelve school districts volunteered to "pilot" the proposed information-evaluation model (Mussio, 1987). The basic intent was to produce a report to each school board describing the achievement of the four goals and six attributes. The report was to include key indicators (from public surveys and Ministry Information Profiles), a summary of district

strengths and weaknesses, and an improvement plan. The project also included an attempt to identify indicators for various categories of Special Education (Mussio, 1987). Five of the districts included this as a subproject.

The project proceeded throughout the 1986-87 school year, with districts distributing the questionnaire component in the spring of 1987. Some difficulties were experienced at this time due to the political unrest in the system. The BCTF, amidst the throes of its ongoing dispute with the B.C. government, "was active in opposing this project" (Mussio, 1987, p. 2). One district dropped out of the project altogether (Victoria), and three others (Vancouver Island North, West Vancouver, Kimberley) scaled down their participation, not producing reports. The remainder (The Arrow Lakes, Armstrong, Campbell River, Nechako, Kamloops, Langley, Surrey, Prince Rupert and Richmond) fulfilled the expectations to a greater or lesser degree.

A number of lessons were learned from the pilot. The first was that most participating districts found the exercise useful and valuable (Dickson, 1988a). A second was that the production of reports for either the professional educator or the public was extremely problematic, due to difficulties with interpretation of data, and presentation of information in a readable fashion. A third was that fear of evaluation is a powerful deterrent to the completion of the exercise. This was a product of two factors: (1) the potential use of evaluation information as a mechanism for political control (Stufflebeam & Welch, 1986); and (2) the distrust and fragmentation of relationships between stakeholders in the B.C. system. The latter concern was not as much of an issue in most of the pilot districts as it might have been, possibly due to three reasons: (a) the "multiperspective" nature of the evaluation; (b) the fact that it was designed to measure the efficacy of the district, and not individual teachers or schools; and (c) the participatory nature of the evaluation (Dickson, 1988). A final message--and an ongoing one--was that the framework and key indicators still need development and refinement.

Stage 3: The IPS in 1988

The IPS in 1988 was still comprised of the four-goal and six-attribute framework. It utilizes Ministry Profiles (see Appendix A2) to provide hard data on the key indicators for each goal, or attribute; and questionnaires to solicit soft data. Questionnaires were designed for students, parents, nonparent taxpayers, employers, ex-graduates, and teachers (see Appendix A3). In addition, assistance with the interpretation of data and reporting out was also provided (Dickson, 1988b). To "model" effective reporting out mechanisms, separate sample reports, one each for the school board and the public, were designed in partnership with the Nechako School District (Nechako, 1987a; Nechako, 1987b). These, along with samples of the hard data available at the Ministry, were distributed throughout the province in a "sample profile".

The Need for Further Refinement.

At the conclusion of the pilot project, the following suggestions were made for future directions:

1. that provincial policies on evaluation and accountability be clarified for schools, districts, and the Ministry;
2. that all schools and districts prepare and file annual reports based on the four-goal and six-attribute framework;
3. that the goals and attributes along with appropriate amendments be endorsed and made "official" by government;
4. that the Ministry, using feedback from the pilots, refine the key indicators and issue easily read and utilitarian information packages, based on the goals and attributes, to all districts and schools;
5. that key indicators be designed in collaboration with other Ministries to provide a common data base;
6. that the questionnaires used in the pilots be revised and made available to schools and districts in the same way the classroom achievement tests are made available; and
7. that the Ministry continue to provide leadership to districts in the interpretation and reporting of data to the public and in the development of improvement plans (Mussio, 1987).

The Spring, 1988, Field Test

While action on these suggestions was considered, a third stage was then generated. Seven school districts further field tested the instrument and provided practical opportunities to address items (4) to (7) above (Dickson, 1988b). Four questions in particular guided the efforts to improve the methodology of the System.

1. How acceptable is the framework of four goals and six attributes in characterizing the overall purposes and methods of the education system?
2. What indicators, other than those currently in use, are appropriate measures for each goal and attribute?
3. What changes should be made to the questionnaires and reporting out formats to make the Information Profile System more effective?
4. What processes and procedures were valuable in focusing evaluation team members on the task of evaluation?

Summary

Political elements in the British Columbia education system have provided an impetus for educational administrators to develop a "leader-substitute" evaluation process. This System, known as the Information Profile System, is designed to assist administrators in dealing with problems of system improvement, motivation and gain of employee and public support, and accountability. Success in meeting of these goals will be revealed in the development of more advanced planning processes, greater employee and public involvement and commitment to education, and an improved understanding of education within the local community.

However, the IPS cannot yet be considered complete. Improvement of its methodology, particularly with respect to the development of key indicators of educational performance, questionnaire data, and report writing is necessary. In addition efforts must be made to determine its effectiveness in meeting the "leader-substitute" demands being made upon it. These are the purposes of this study. The next step in achieving these purposes is to examine both the literature on administration and evaluation to seek out strategies and

principles that would enhance the general theoretic and practical framework of the IPS model.

Part II: The Literature on Administration

The IPS is a holistic evaluation model that is designed to assist administrators to accomplish the goals of improved accountability, gain of public support, and improvement of school district achievement. Prior to its use it may well be considered a holistic evaluation model, but during its use--even in action research as employed in this dissertation--it is also a leadership initiative. The IPS can be implemented only through a series of administrative acts and behaviors. Thus, as well as providing insight and understanding into the concept of evaluation, observation and involvement in the use of the IPS will also provide insight into the concepts of leadership and administration. Even more important is the development of a better understanding as to how evaluation can fortify administration, and vice-versa.

The Nature of Knowledge and Theory in Educational Administration

If one is to lay a knowledge base for a dissertation in educational administration the first step is to deal with two main questions: (1) what is the nature of knowledge in general; and (2) how is a knowledge pool for a particular topic such as educational administration best constructed?

The answers to these two questions have been the source of great debate during the last twenty years. Two main camps have developed. The camp of orthodoxy believes in a "science" of administration and has, as major proponents such writers as Willower, Hills, and Griffiths (Gronn, 1983). The challenging paradigm, best termed a "subjectivist" viewpoint (Gronn, 1983) has its main proponent T. B. Greenfield. His main contention is that there can be no single theory of administration due to subjective nature of reality and of how one comes to know reality. A more detailed examination of the salient viewpoints of each perspective will now be presented.

Philosophical Premises of the Emerging Paradigm

Paradigms are basic world-views, or belief systems through which one orders experience to create models or patterns that enhance understanding. These world-views are fundamentally philosophical, and one can only be understood in contrast with the other; in a sense one defines the other.

The emerging paradigm, or subjectivist viewpoint, has seven main assumptions that distinguish it from the more popular scientific one. These are summarized from Lincoln (1985) :

1. A shift from a belief in a simple and probabilistic world (scientific paradigm) to a belief in a complex and diverse world (subjectivist paradigm).

2. A shift from a belief in a hierarchically ordered world, i.e., a world with well-ordered, sequential levels of "correct" theory, leading one to ultimate truth (scientific paradigm); to a belief in a heterarchical one, in which there may be many different knowledges and theories, all of which are correct in light of the individual context in which they were developed. In the subjectivist paradigm, there is no truth, only understanding.

3. A shift from a belief in the image of a mechanistic and sequentially ordered universe (scientific paradigm) to a belief in a holographic one (subjectivist paradigm). The latter suggests that the components of a whole can be both independent and interdependent, with each part containing the essence of the whole within it.

4. A shift from the image of a determinate universe (scientific paradigm) to an indeterminate one (subjectivist paradigm). Science is predicated on objectivity leading to definitions of reality and culminating in predictability; indeed this is the main point, or purpose of scientific methodology. In the subjectivist paradigm there is no belief that the world can be predicted or controlled; the future is ambiguous, and indeed is constructed through choice, even choice of methodology itself.

5. A shift from an assumption in direct causality (scientific paradigm) to mutual causality (subjectivist paradigm). In the former case, one is predisposed to believe that the concurrence of one event is directly attributable to a preceding one; in the latter case, the belief is that linearity of causality is not necessarily so. Indeed, the existence of some act can in fact be both the

cause of as well as the result of another act: they are mutually dependent upon each other for their existence.

6. A shift from a belief in the metaphor of assembly--ie., the whole being equal to the sum of the added parts (scientific paradigm)--to a belief in morphogenesis (subjectivist paradigm). Morphogenesis refers to the belief that a whole can be more than the sum of its parts; in joining together segments, or units, a whole new completely different entity can be created. Knowledge of that entity can not be gained by analysis of its constituent pieces.

7. a shift from a belief in pure objectivity (scientific paradigm) to one of subjectivity (emerging paradigm). In the former case there is a belief in independent reality, and the ability to abstract from knowledge of it to fundamental truth. In the latter case, knowledge of reality is perspectival--dependent on one's own interaction through the mind with that object, independent reality or not. Multiple realities can then be constructed around the same phenomenon. Ultimately research becomes the task of collecting, collating, and understanding these multi-perspectives so one can make sense of the world.

A belief in the emerging paradigm suggests that the mechanistic view of organizations, of administration in organizations, and of evaluation be reconstituted so as to reflect the above principles. The literature review in the remainder of the chapter presents material apropos to these suggestions. It also suggests a naturalistic approach to the design of the study, for reasons that will be discussed more at length later in the chapter.

The Impact of the Emerging Paradigm on Educational Administration

Most studies in educational administration in the past fifty years have been constructed on the supposition that knowledge of human action in educational organizations can be understood through the same methodological approaches as are applied to gain understanding of the material world in the physical sciences. Griffiths (1983) states for example that "virtually all of the research being done in educational administration can be characterized by what...(is called) "normal science", that is, research on problems derived from the now-accepted paradigm" (p. 119). Implicit in this belief is acceptance of the fact that human behaviors can be seen as facts, that in observing and generalizing on these facts one can build theories. These theories can then be tested

deductively and if seen to be true will become "laws" of human behavior. In the scientific mode of inductive-deductive, rational investigation-proof, one is led towards an immutable truth, and the certitude of predictability.

Willower (1983) vehemently argues that reliance on the scientific paradigm has served society well both in the physical sciences as well as the social sciences. He views the debate as a philosophical one (i.e., the phenomenologist viewpoint versus the empirical, positivist one) but indicates that in a pragmatic sense these issues are relatively unimportant. Whatever one views the nature of reality to be is immaterial because even the scientist can only deal with that reality as it is revealed to us in our interactions in the world. Not accepting subjectivist epistemological arguments, Willower rejects the concept of science as an ideology, claiming it is in fact simply a commitment to a particular mode of inquiry, not to a wider set of beliefs or principles. He embraces qualitative methods as appropriate social science techniques. He also claims that the scientific viewpoint is not incompatible with understanding, even empathetic understanding, in its pursuit of knowledge.

Griffiths (1983) provides a slightly different viewpoint. He argues that much good work has been done using the scientific method, stating that the scientific paradigm need not be abandoned as a source of investigative inquiry. However, he then goes on to state that the application of scientific methods to administration is enough in question so as to encourage us to "develop much more restricted theories" (p. 134). Claiming that transfer of knowledge from one context to another is the point of research, Griffiths outlines a set of criteria that enhance the potential for transference of knowledge from one organizational context to another. His opinion is that theoretical knowledge developed through the scientific method is acceptable when the organizations studied are much alike.

The leading proponent of the subjectivist field of thought in educational administration is T.B. Greenfield (1975, 1979, 1980). In presenting a number of arguments based upon the emerging paradigm, (see Table 2.1) he makes two major claims that are of importance here: (1) that no single theory of administration can exist, because theories as constructed in science are based on the concepts of reality and objectivity which the subjectivist paradigm demonstrates are not epistemologically sound; and (2) that knowledge gained

about organizations, and administrative behavior in organizations is best derived through methodological principles emphasizing the phenomenological perspective.

In keeping with these two viewpoints, he suggests a number of prolegomena to be used as guidelines in the study of administrative behavior in organizations. The key propositions are presented below:

Proposition 1: That organizations are accomplished by people and people are responsible for what goes on in them.

Proposition 2: That organizations are expressions of will, intention and value.

Proposition 4: That facts do not exist except as they are called into existence by human action and interest.

Proposition 5: That man acts and will judge action.

Proposition 6: That organizations are arbitrary definitions of reality woven in symbols and expressed in language.

Proposition 7: That organizations expressed as contexts for human action can be resolved into meaning, moral order, and power.

Proposition 9: There is no way of training administrators other than by giving them some apocalyptic or transcendental vision of the universe and of their life on earth (Greenfield, 1975, p. 100-107).

These propositions focus on understanding organizations as human constructs, as opposed to legal or technical realities. The study of educational administration then becomes a study of human interactions within schools or school districts. These interactions are better revealed through analysis of values and emotions than through analysis of the organization's structural components. Knowledge is a function of individual experience within a particular context, and is gained through personal engagement as opposed to impartial observation. Knowledge is transferable when understanding derived from one context is applicable in the person's own judgement, to another context.

Table 2.1: Greenfield's analysis of social reality. Taken from Gronn, (1983), p. 68-9.

<i>Dimensions of comparison</i>	<i>What is social reality?</i>	
	<i>A natural system</i>	<i>Human invention</i>
Philosophical basis	Realism: the world exists and is knowable as it really is. Organizations are real entities with a life of their own.	Idealism: the world exists but different people construe it in very different ways. Organizations are invented social reality.
The role of social science	Discovering the universal laws of society and human conduct within it.	Discovering how different people interpret the world in which they live.
Basic units of social reality	The collectivity: society or organizations.	Individuals acting singly or together.
Method of understanding	Identifying conditions or relationships which permit the collectivity to exist. Conceiving what these conditions and relationships are.	Interpretation of the subjective meanings which individuals place upon their action. Discovering the subjective rules for such action.
Theory	A rational edifice built by scientists to explain human behaviour.	Sets of meanings which people use to make sense of their world and behaviour within it.
Research	Experimental or quasi-experimental validation of theory.	The search for meaningful relationships and the discovery of their consequences for action.
Methodology	Abstraction of reality, especially through mathematical models and quantitative analysis.	The representation of reality for purposes of comparison. Analysis of language and meaning.
Society	Ordered. Governed by a uniform set of values and made possible only by those values.	Conflicted. Governed by the values of people with access to power.
Organizations	Goal oriented. Independent of people. Instruments of order in society serving both society and the individual.	Dependent upon people and their goals. Instruments of power which some people control and can use to attain ends which seem good to them.
Organizational pathologies	Organizations get out of kilter with social values and individual needs.	Given diverse human ends, there is always conflict among people acting to pursue them.
Prescription for curing organizational ills	Change the structure of the organization to meet social values and individual needs.	Find out what values are embodied in organizational action and whose they are. Change the people or change their values if you can.

The best training for educational administrators is then to engage in experiences that allow them to gain larger and more indepth understandings of the nature of life in educational organizations.

Applying the Emergent Paradigm to This Dissertation

In designing a dissertation study one must adopt a philosophical stance that reflects one's own philosophical and epistemological views about knowledge, and about research in educational administration. Having been a practising administrator for thirteen years I am persuaded that Greenfield's viewpoint describes the approach most conducive to gaining meaningful knowledge for one's own personal growth. If such knowledge is constructed according to methodologies related to the subjectivist viewpoint (i.e., through naturalistic inquiry) there is also the possibility that it can be transferred by others to their own context if they deem circumstances to be comparable.

Action research is a procedure consistent with this philosophical stance. For this reason it was employed as a procedure in this study. However, thirteen years of field practice has also validated other aspects of administrative theory. For this reason the next section of the chapter--a review of literature on leadership/administration--is included, as many of the concepts so contained recommend themselves to a better understanding of administration-in-action. Part IV of the literature review shows how the subjectivist paradigm impacts on the second forms of the inquiry: evaluation. As my knowledge in this area has only been slightly informed through practice, a broader theoretical overview is provided so that when engaging in the action research I will be able to identify those elements apropos to the real-life situation.

The design of the action research is presented in Chapter III and the resulting case studies in subsequent chapters. The culmination of this activity will be, of course, the development of a process for district evaluation that will be informed by both practice and theory, and that can be used if administrators see it as applicable to their personal situational context.

Part III: Leadership/Administration Theory and Organization Theory

Evaluation can be used as an administrative tool. In order to be most effectively utilized it must be developed and employed in light of effective administrative knowledge and action. In order to understand evaluation's role in administration, and its potential as an effective leader-substitute, the genesis of this theory and related information about effective leadership in education organizations need to be discussed.

The Leader-substitute Construct

Many studies have been done on the characteristics of successful administrators. Yet leaders who function well in one context often do poorly in another (Doig & Hargrove, 1987). The leader-substitute theory suggests that leadership (i.e., the capacity to lead) not only lies with the formal leader in the organization, but also with other characteristics of that organization that exist independent of and only within his peripheral control (Kerr & Jermier, 1978). It is the interaction between these characteristics of the organization (often unknown or consciously perceived even by the organization's own members) that makes administration (the whole) different from the abilities of the administrators themselves. The leader-substitute theory provides an explanation as to why research into leader behaviors, styles, and situational response yields so few consistencies (Freeston, 1986).

A fundamental cornerstone of the leader-substitute construct is that power--the ability to influence which is key to leadership--can lie outside people: acting on people, but not attributable to people, in that their immediate efforts to influence, by dint of personality, character, or behavior, appear to be counteracted by larger forces that are culture based, or structure based. In this sense the leader-substitute concept of leadership differs from classical conceptions, in which personal power, embodied in expertise, reward, status, or charisma (Etzioni, 1980) is the acknowledged source of influence in organizations.

The Role of Power and Leader-substitutes.

Hiley (1987) discusses the role of power in administration. He describes three forms of power that provide the cornerstone of administrative action. The first--"person-based" power--is the principle upon which most traditional theories of leadership are based. The second and third--"ideology-based" and "disciplinary-based" power--provide a basis for justification of the leader-substitute theory.

Hiley (1987) suggests that in traditional conceptions of leadership, such as trait theory, contingency theory, and attribute theory, power--the ability to influence and therefore lead--is a feature of the relation between individuals. Differing levels of "person-power" creates leadership and followership in a relationship. The trait theory, the attribute theory, and the contingency theory of leadership are predicated upon this principle.

When someone says "That person is a natural leader"--they are using the term to describe a quality that person can call upon to produce the conditions necessary for leadership. This use of the term "leader" suggests that leadership is a trait, explainable through a study of the characteristics and actions of leaders. Indeed, trait theory "dominates the research in educational administration" (McPherson, Crowson & Pitner, 1986, p. 225).

From trait theory have come various characterizations of leadership. Three of the best known are Halpin's Leader Behavior Description Questionnaire (LBDQ), Likert's work at the University of Michigan, and Blake and Mouton's "Managerial Grid" (McPherson, Crowson, & Pitner, 1986 p. 229-31). In each case the belief is that the existence of the trait of leadership is revealed through measurable behavioral actions exhibited by the leader. All three approaches concentrate on two bipolar dimensions of leadership: the idiographic (people oriented dimension), and the nomothetic (institutional dimension) (Getzels & Guba, 1957). Thus, Halpin talks about the consideration versus structure-initiating functions of leadership; Likert about the need for "employee-centered" leadership versus "job-centered"; and Blake and Mouton¹

¹ Blake and Mouton, in constructing their "managerial" grid, used the term "administration" to describe the various dimensions of leadership. They talk about Style 1,1--Caretaker Administration, and Style 1,9--Comfortable and Pleasant Administration (McPherson et al, p. 230). The semantic confusion here is increased because the term "managerial" is identified

about the concern for production versus the concern for people (McPherson, Crowson & Pitner, 1986, p. 226-32).

Attribution theory--the concept that leadership is a function of followers granting to leaders the power to lead--also stresses the "person-based condition" as a fundamental tenet. This theory states that:

Followers grant an individual the obeisance due a leader when it is possible for them to explain (attribute reasons for) the behavior of that leader in terms that are consistent with their own preconceptions. (McPherson, Crowson & Pitner, 1986, p. 245)

In other words, if followers perceive that a person appointed to a position of formal leadership (i.e., principal of a school) has the right to influence them in the area of school policy and procedure, then they will "obey" when a directive concerning attendance procedures is forthcoming. Simon (1976) calls this the "zone of acceptance"; the degree to which employees will accept leader authority without questioning the validity of that authority. This is due to the followers' preconceptions as to his responsibility, or "right to lead". If on the other hand, a directive is issued, say, defining appropriate dress for teachers, and appropriate dress is seen to be outside the realm of responsibility of the principal, it might well not be obeyed: thus, the zone of acceptance has been breached and leadership is not exercised. Power in this sense is a function of a dynamic relationship between leader and follower, and remains person-based.

The situational theory of leadership suggests that leadership effectiveness is only partly a product of the personal qualities, or behaviors that a formal leader exhibits. The context within which the leader is operating can either enhance or mitigate the effect of leader behavior (Doig & Hargrove, 1987). For example, Fielder's contingency theory of leadership argues that leadership performance "is dependent upon the interaction of leadership style and the favorableness of the situation" (McPherson et al, p. 234). He goes on to describe leadership effectiveness as a function of two kinds of behaviors (consideration and task-orientation) interacting with environmental conditions: one style is more preferable in certain circumstances, and the

with both leadership and administration. The relationship between the three will be clarified later in the paper.

other in different circumstances. Again, however, the focus is on leader behavior--as a function of innate leadership ability--and therefore equates leadership (the act) with leadership (the quality).

Inherent in contingency theory, as in the "trait" theory, is the belief that the leadership quality is inherent within an individual; it is a function exercised solely by the human condition--albeit constrained or enhanced by the context within which it operates. Therefore it is an act only observable as manifested through a person. In other words, this theory--as well as the "trait" theory--does not conceive of the act of leadership existing (by that we mean the conditions ascribed to leadership being in evidence) without human initiation: that is, without a person employing that quality also called "leadership". As the term is used in classical leadership theory, leadership is a function of the human condition--its existence is dependent upon power being exchanged between people.

"Ideological" Power and "Disciplinary" Power

Hiley (1987) suggests, however, that there are two alternative sources of power. The first is "ideological power", whereby the ability to influence is located in a generalized system of beliefs or values. It is this conception of power that is consistent with the theme found in much leadership literature today; i.e., the writings of Deal & Kennedy (1982), Johnston (1987), Sergiovanni & Corbally (1984), and Stout (1987), stressing the impact the culture of the organization has on leader effectiveness. The myths, rituals and ceremonies of the culture of the organization "provide a sense of common direction for all employees and guidelines for their day-to-day behavior" (Deal & Kennedy, 1982, p. 21). This statement implies that the "informal" ideology of an organization exercises leadership--as much as the style or behavior of the formal leader does. So ideological expectations of task, and cultural attachments of subordinates will act as surrogate influencers according to this theory.

Hiley's (1987) third conception of power is "disciplinary power". He revisits the writings of Michel Foucault, a French philosopher and historian, who suggests that whereas person-based power can be repressive, disciplinary power is not:

...unlike a commodity exchanged between individuals, it is continually operative throughout a social structure like a field of force...it operates by organizing our bodies in time and space and its goal is to achieve maximum utility while simultaneously achieving maximum obedience. (Hiley, 1987, p. 351).

It is this form of power that is embodied in formalized structure, such as role descriptions, policies and procedures, and regulations. Hiley (1987) also gives the example of assembly line speed-ups to coerce workers to be more productive (p. 352).

These three forms of power--person-based, ideology--based, and structure-based--support the existence of "substitutes for leadership". The extent that subordinates within an organization are influenced by their own innate abilities and predispositions or by the ideological interactions within that organization will determine their response as followers; indeed, they themselves might be the "leadership" component within that organization. Similarly, structural factors relating to issues of task and procedure exercising "disciplinary power" can and will act as "leader-substitutes". So too will those aspects of task embodied in the "culture" of the organization.

The Leader-Substitute Construct

Some authors (Ford, 1981; McPherson, Crowson & Pitner, 1986) see the leader-substitute theory as a form of the contingency approach to leadership. As such, the theory claims to identify those salient aspects of the situation "that influences the consequences of leader behavior" (Ford, 1981, p. 274). Kerr and Jermier (1978) identify three main aspects of the formal organization that act to alter the effectiveness of leader behavior. These are the characteristics of task, the characteristics of subordinates, and the characteristics of structure that exist within an organization. Implicit in their argument is that cultural traits of an organization (i.e., the existing patterns of human interaction), structural attributes of the organization (i.e., degree of organizational formalization, in the form of goal statements, policies, etc.), and the nature of tasks people must perform within the organization (i.e., routine, or methodological tasks versus creative, adaptive tasks) determine the style and behavior required of the "status" executive in order to demonstrate effective leadership. Hodgkinson (1983) lends support to their empirical investigation through his treatment of leadership from a philosophical point

of view; he states that leaders have four responsibilities: (1) know the task; (2) know the situation (structure); (3) know his followership; and (4) know himself.

Kerr and Jermier (1978) go so far as to suggest that these leader-substitutes can replace or 'act in place of' a specific leader behavior, and can "render hierarchical leadership both unnecessary and impossible in terms of the potential impact of leadership on important subordinate outcomes" (Howell & Dorfman, 1981, p. 715). Freeston (1987) extends this argument, claiming that the qualities of consideration traditionally attributed to "person-leaders", such as guidance and good feelings, can in fact derive from characteristics of the individual, the task, or the organization--not from leader behaviors:

Called "substitutes for leadership," this construct holds that some characteristics may have a greater effect on subordinate satisfaction and performance than do leader behaviors. *The leader substitutes construct rejects the assumption that hierarchical leadership is always important to subordinate outcomes.* (Freeston, 1987, p. 45; emphasis added)

Gamoran and Dreeben (1986) claim, for example, that in "loosely-coupled" educational organizations (Weick, 1985; March & Olsen, 1976), the locus of control--that is, the ability to influence teaching outcomes--lies primarily in resource allocation. To the extent that teaching time, textbook purchase, and class-size is determined by formula, they argue that the formula is the influencing factor; the hierarchical leader is not the meaningful component here in providing leadership.

The Leader-Substitute Construct and Administration

The existence of leader-substitutes suggests an explanation as to why the terms "administration" and "leadership" are often confused. The following statements illustrate this:

Administration is leadership. Leadership *is* administration. These identities need surprise no reader who has come this far *but still there is a certain strangeness to these expressions which is a carry-over from common usage. There is a tendency in the ordinary language to conceive of leadership, loosely understood, and it is nearly always very loosely understood, as if it were sort of an increment to the administrative-management process which might or not be present. An epiphenomenon.* As if one could administer without leadership or lead

without administration...it somehow does not make good sense to talk of having at one and the same time good leadership and bad administration, or conversely. (Hodgkinson, 1983, p. 195; emphasis added)

and:

Although there seems to be considerable disagreement over the relationship between leadership and administration, *most would agree that there is no necessary relationship between these two types of activities. It is logically impossible to engage in administration without exercising leadership...* (Coombs, 1988, p. 11; emphasis added).

Coombs (1988) sees the concepts of administration and leadership as independent because in her mind, leadership is "person-based", and therefore only exercised as part of the administration function within a school. Hodgkinson, on the other hand, sees leadership as more than person-based:

...leadership can be understood as the effecting of policy, values, philosophy through collective organizational action....To phrase it differently, *either leadership or administration is the moving of men towards goals through a system of organization. It can be done well, or it can be done badly, or it can be done indifferently, but it cannot not be done at all.* (Hodgkinson, 1983, p. 196; emphasis added).

Administration is both "person-based" and "procedure-based". Thus when the disciplinary influence of day-to-day routines takes over directing operations of the organization, administration is still acting, whereas "person-based leadership" is not. All the necessary and sufficient conditions for leadership are being met within the organization, but they are being provided by a nonhuman element. This characteristic explains how administration can be constant in operation, and "omnipresent". The quality of leadership is exercised by the processes and procedures developed as part of administrative policy (See Figure 2.1).

Dunsire (1971) states that a major component of administration is the concept of service. Whereby "'leadership' is usually viewed as unsettling the status quo, and 'management', as keeping things running smoothly, 'administration' seesaws in meaning from one to the other of these two terms" (Coombs, 1988, p. 10). Organizations and organizational members need the regular, omnipresent assistance that administrative support staff, management policies and cultural traditions provide. Service, as well as

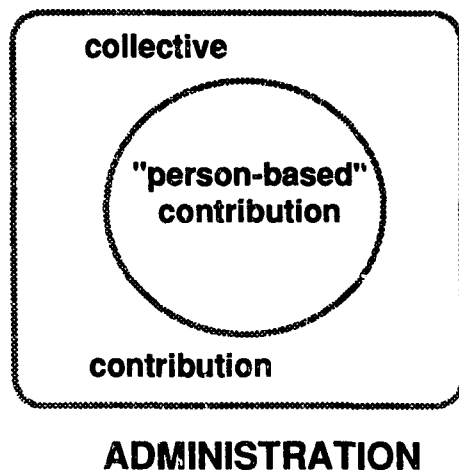


Figure 2.1: The two components of administration.

leadership, is integral to the word "administration". This service is usually provided by the policies, procedures and support (i.e., clerical) that exist within that organization, and the "person-leader" is called upon for his assistance only in circumstances where they do not suffice.

Effective Administration in Loosely Coupled Organizations

Educational organizations do not fit the traditional picture of rational, specialized, line authority bureaucratic organizations as described by Weber (Clarke, 1985). And as the educational organization unit gets larger, i.e., school district levels of organization as opposed to the classroom level, the fit is even less justifiable. March calls education systems "organized anarchies"--lightly controlled, unstable, filled with ambiguity, segmented with a great deal of discretion and control within each segment, and little controlled by rules, policies, and procedures (McPherson, Crowson & Pitner, p. 95).

Weick (1985) describes the loosely-coupled organization concept as resulting from "new paradigms of inquiry" (p. 106). He depicts them as less dependent on rationality than upon tradition, personal value, and informal operating habits. Such organizations are usually broken into small independent segments in which there is relative autonomy of action and order contained within (i.e., schools, and teacher classrooms within schools). They are also subject to great ambiguity in terms of goals and purposes. This is due

to conflicting interests and desires by different groups (such as professional teacher's associations, school administrator groups, school boards, etc.); as well as the individual desires each person brings to the organization .

Floden, Porter, Alford, Freeman, Irwin, Schmidt and Schwille (1988) examined a large number of school districts to see if this image was reflected in the practice of instructional leadership of math teaching. Evidence indicated that control, (i.e., tight coupling) was a function of three characteristics: consistency, prescriptiveness, and supporting policies with authority. Consistency and prescriptiveness looked at whether teachers were provided with conflicting expectations re textbooks and required content within those textbooks. Supporting policies with authority was described as providing teachers with good reasons based on expertise as to why the guidelines should be followed. While appearing rather tightly coupled in terms of consistency and prescriptiveness, most districts did not provide the leadership in exercise of either authority or power to support educational policies. As such they retained the loosely coupled characteristic.

Effective leadership is described by Floden et al (1988) as "...persuad(ing) teachers that there are good reasons for the district policies", and developing "trust and belief in the personal or group that has established the policy" (p. 115). This definition of effective leadership is reflective of the new attitude to administration developing in response to acceptance of the concept of loosely coupled organizations; it suggests flattening of the hierarchical leadership practised in traditional bureaucratic organizations.

As long ago as the early 1970's there was intellectual acceptance of the need for a paradigm shift in leadership, moving away from the traditional, authoritative, bureaucratic style of leadership to a less traditional, democratic and shared approach. Yet conversion of the intellectual to the practical operations of school administration is still a recent phenomenon:

Schooling in America today is characterized by a new politics of education and a fundamental shift in the basic governance structure for organizing and administering schools. Much of what has happened and is now happening is only now being sorted out, and the implications for administrative practice in the decades ahead are only now beginning to be understood (Sergiovanni, Burlingame, Coombs & Thurston, 1987, p. 1).

Principles of Effective Administration

This section of the study is devoted to explaining some of the new basic principles of effective administration, suggesting that they be considered when developing administrative tasks and structures, and when acting through subordinate characteristics. If applied to school district administration, these principles should maximize the likelihood that the IPS will be an effective administrative tool.

Administration, or "Leadership has two bases: a moral component and a technical (one)" (Stout, 1986, p. 199). This statement rephrases the normative and ideographic dimensions of leadership (Getzels and Guba, 1956), and the consideration and structure-building dimensions described by Halpin and Winer (1956).

Effective administration emphasizes the moral component of leadership. Barnard (1938) in his seminal book The Functions of the Executive stressed this point when discussing executive responsibility: The moral component is:

...the aspect of leadership we commonly imply in the word "responsibility", the quality which gives dependability and determination to human conduct, and foresight and ideality to purpose (Barnard, 1938, p. 260).

In Barnard's estimation the true art of leadership was maintaining a moral balance between one's responsibility and one's role. With responsibility come obligations, and as one rises into the upper echelons of the organization the complexity of both those responsibilities and obligations increases. This places an increasingly greater strain on the executive's ability to reconcile conflicting moral viewpoints--for as one shifts from the floor of the factory to the head office, or from the playground to the school board office, decisions tend to become more philosophical--and of greater moral import--than they are practical. Thus the strain of executive life is having to constantly resolve moral dilemmas. If however, the executive does not have the character, or the fortitude that these issues require for their successful resolution, then that executive has responsibility but not commensurate moral leadership ability. Barnard puts it this way:

...morals are personal forces or propensities of a general and stable character in individuals which tend to inhibit, control, or modify inconsistent immediate specific desires, impulses or interests, and to intensify those which are consistent with such propensities....When the tendency is strong and stable there exists a condition of responsibility (p. 261)

Responsibility is a function of two things: (1) the strength of the individual's personal moral code; and (2) the demands made upon that moral code. As one moves up the levels of a hierarchy, into positions of greater responsibility, the demands made upon one's moral code are greater; if one cannot bring the character and principles of the demanded morality to bear, then he is not fulfilling those responsibilities for the benefit of the organization.

Barnard (1938) also makes the point that moral principles are derived from without the individual rather than solely from introspection. Morality arises at the interface of the individual and society; it is a set of guiding principles established to preserve agreed upon values, or rights, that are threatened through improper interaction. However, these values are also affected by organizational context: As C.S. Lewis (1943) states, practical action based on moral acts is situationally determined. Thus, in public service organizations such as education "giving dependability and determination to human conduct" means that the administration must emphasize those values and purposes that are fundamental to society, and to education (Coombs, C. 1988).

The Role of Values in Moral Administration

The emphasis on morality requires an understanding of the role values play in administration. Holmes (1986) brings the issue of values and their role in educational administration into clearer focus. He states that:

The lack of consensus about the purpose of the elementary and secondary school makes it more important rather than less to have a clear framework of goals and values. The modern idea that schools can function in a value-free atmosphere brings the whole educational profession, and particularly administrators, into disrepute (p. 80).

Weick (1985), in discussing loosely coupled organizations, explains that greater ambiguity--a quality of education systems--"increases the extent to which action is guided by values and ideology" (p. 125). He goes on to state that

people who can resolve negative ambiguity (there is a dynamic tension here; some ambiguity is necessary and desirable in a professional context) can gain power. In situations of negative ambiguity administrators and organizations can seize the opportunity to reshape the vision of the organization and its role in society. They can do so by altering existing ideologies by the imposition of a new set of values for the organization. Weick (1985) states:

When new values are introduced into an organization, a new set of relevancies and competencies are created that can provide a badly needed source of innovation...In this way ambiguity can produce innovation and greater utilization of resources (p. 125).

Sergiovanni, Burlingame, Coombs & Thurston (1986) support this viewpoint when they discuss the need for administrators to develop a framework of immutable values within an organization and to ensure that all decisions are decided in light of that value framework. Thus the technical components of leadership, i.e., the policies and procedures of the organization, must be constructed so as to ensure that relationships can be predicated on principles of openness, rationality, justice, equal consideration, truthfulness, and reciprocal obligations (Wright, 1986). In addition, organizational values of quality, efficiency, and effectiveness must also be promoted (Barnard, 1938). Moral leadership in an organization stresses the maintenance of the organization's value *raison d'être*, while at the same time preserving values of human dignity.

Administrators operate in a continuum from philosophy--establishing the basic values of an organization--to practice, as expressed through management action: Controlling and supervising the day-to-day "use of resources to achieve intended outcomes" (Wholey, 1983, p. 10). Moral leadership is administration that maintains a productive tension between philosophy and action (Zuckert, 1988). Effective leaders employ philosophy to curb the natural tendency of organizational members to overemphasize individual "spiritedness", or emphasis on self:

Executive responsibility, then, is that capacity of leaders by which, reflecting attitudes, ideals, hopes, derived largely from without themselves, they are compelled to bind the wills of men to the accomplishment *of purposes beyond their immediate ends, beyond their times.* (Barnard, 1938, p.283; emphasis added).

Philosophy as expressed in a basic values statement is a counterfoil to action, as it places immediate events and issues in a larger context. Moral, and effective leadership finds the balance between the two, allowing the basic values pertaining to the organization to direct it (Johnston, 1986) while at the same time allowing individual discretion and action commensurate with people's need for self-identity and esteem.

Hodgkinson (1978) suggests that the philosophical challenge of administration is to sort out value priority in the organization, and for effective leaders to emphasize "right" versus "good" values. He defines values as "concepts of the desirable", and suggests three levels of value priority (see Figure 2.2)

Value Type	Grounds of Value	Psychological Facility	Philosophical Orientations	Value Level
I	PRINCIPLES	conation willing	religion existentialism intuition	I
IIA	CONSEQUENCE (A)	cognition reason thinking	utilitarianism pragmatism humanism democratic liberalism	II
IIB	CONSENSUS (B)			
III	PREFERENCE	affect emotion feeling	behaviourism positivism hedonism	III


RIGHT

 GOOD

Figure 2.2: The Value Paradigm (taken from Hodgkinson, 1983, p. 38).

The highest level consists of principles derived from a religious calling or ideological commitment. The middle level is comprised of moral or consensual values (i.e., those principles accepted by society as fundamental to its maintenance, or agreed upon for their desirability). Preferential values--those of an immediate and compelling nature that often conflict with higher level values--are the lowest strata. Hodgkinson stresses the need for the administrator to extend philosophy into practice by bringing to the job a clearly defined philosophy of education, based on ideological, moral and consensual values; and then finding ways to balance that philosophy in light

of the preferential values that drive the organization on a daily basis. In this fashion value priority can be established so as to resolve value conflicts. Evaluation is a process by which this can be achieved.

Developing a Values Focus

When contemplating action administrators must ensure that the values fundamental to the organization are preserved, and if possible, enhanced. To accomplish this administrators must both lead and serve (Dunsire, 1971). Administrators lead by developing future directions for an organization that reflect organizational values, and binding the wills of men to them; they serve by designing procedures, structures and tasks that respect the values of individual workers and clients. Once again, keeping a balance between the two disparate functions is the moral challenge to leadership.

One way effective administrators achieve this balance is through management of meaning. Management of meaning is a process by which administrators try to focus organizational members on a common language, common information pool, and common value system for the organization. If all members of the system see education in the same light, the potential for acceptance and commitment to improvement plans system-wide is enhanced.

There are three ways by which administrators can manage meaning in loosely-coupled organizations like schools, or district education systems. These are: (1) by designing an empirical "vision" of the organization's future; (2) by developing a sense of teamwork and empowerment through the use of democratic and participative decision-making in leadership activities; and (3) by controlling and managing information flow by being an information-broker.

Vision is the ability of the administrator to picture where the organization should go in the future (Storey, 1988). Implied is a belief that the administrator has knowledge both of the values he wishes to inculcate in the organization, and the values held by other constituents (i.e., teachers, parents, students, and community members). Also implied is that the administrator can build these values into structures, tasks, and procedures that allow for their maintenance in day-to-day activity--sometimes subjugating immediate, preferential values to longer term, moral or rational values, while at other times letting immediate concerns hold sway. In this way

philosophical principles extrapolate into practice--and morality, which requires differing modes of practice in varying situations to achieve the same moral principle, is achieved.

Effective visions are not, however, defined solely by the administrator (Storey, 1988). Successful vision must be communicated to others in the education system. Once communicated, it must be open to change--to inculcate other constituent aspirations into a "empirical vision" for the future. It must be a pragmatic vision, one with an empirical base. A vision that does not reflect realities of current operations, so people can see how to get from "what is" to "what should be", is a vision of little substance to most members of the organization (see Figure 2.3).

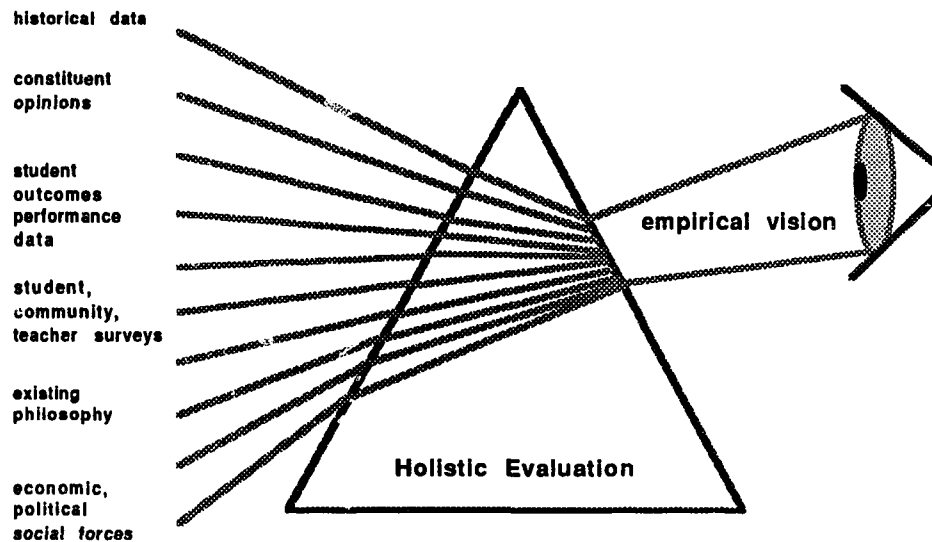


Figure 2.3: Developing an empirical vision (adapted from Storey, 1988).

A second method by which administrators can manage meaning is to enhance other organizational members' leadership capacity. March (1981) makes the comment that "Administrators are vital as a class but not as individuals...What makes an organization function well is the density of administrative

competence... (in that organization)" (p. 29). A mechanism to accomplish this is "transformational leadership" (Burns, 1978):

Transformational leadership is more concerned with end-values such as liberty, justice, equality (p. 426).

That people can be lifted into their better selves is the secret of transforming leadership and the moral and practical theme of this work (p. 462).

effective administration employs procedures by which the capacity to lead is disseminated throughout the organization, and similar to Barnard's concept of responsibility, people are empowered (i.e., provided with the understanding or shared purpose, a sense of responsibility for that purpose, and a feeling of confidence in their ability to achieve it) to demonstrate leadership in the pursuit of their organizational responsibilities. Morality is enhanced when people are held accountable for achievement of their delegated responsibilities.

But how can empowerment be achieved within an organization? Morgan (1987) suggests an answer. A first step is to develop a metaphorical image of the organization, one that reflects the basic qualities or purposes of its *raison d'etre*. One such image particularly appropriate to education administration is the image of "organization as brain." Since loosely coupled organizations are systems within systems, one can view that segment of the organization imbued with the responsibility of administration as an organization in itself; an organization that should act as the "brain" of the overall system.

Two qualities possessed by this image are (1) its inherent "learning" capacity; and (2) its holographic nature. These two ideas suggest that educational administrators encourage within each person in that organization the same problem-solving capacity that one wishes the administration of the organization itself to adopt; and then devise parallel structures, policies, and procedures organization-wide. Morgan (1987) even goes so far as to suggest that evaluation is, in fact the brain's natural learning process. Devising evaluation procedures for an organization as a whole should infuse it with a problem-solving capacity, and if this capacity is mirrored by the individual actions taken by each member, commensurate with their responsibility in the

organization, one can assume that empowerment will result (Morgan's suggested methods are described later in this chapter).

Another suggested method to accomplish empowerment is to democratize the organization so that each individual should be allowed input into decision making (Bates, 1985). He suggests that effective administration inculcates four desirable changes: "The democratization of social relations, the democratization of knowledge, the democratization of communication, and the democratization of cultural concerns" (p. 98). Collegial leadership is designed to overcome traditional bureaucratic impediments leading to fragmentation, and to develop a sense of team and shared responsibility. Participatory, democratic leadership rather than authoritative, hierarchical leadership allows administrators to overcome the tendency of an organization to fragment in spite of the commonality of purpose, and infuse it with the ability to unite as a team in spite of the differences dividing the organizational members.

Bloom (1987) cautions against the abuse of democratic principles, however. If these principles are conceived from a basic belief in natural rights, and the proper balance between natural rights and individual action is maintained, then democratic principles can guide the organization toward a moral focus and sense of empowerment. If, on the other hand, values of liberty and equality are overemphasized, as is the penchant today, then openness and relativism results and morality is lost. Openness with relativism is denigrating to the true educational spirit; openness with reason ennobles it. Bloom (1987) argues, however, that the tendency today is to rely too much on scientific methodology and not enough on our natural spiritualism. This leads to an overemphasis on empirical evidence interpreted in a context of "no right and wrong". In constructing his material world based on scientific principle man has removed himself so far from nature as to lose contact with natural rights and values. Thus the productive tension between philosophy and practice has been disturbed. In education, at the philosophical end of the continuum are the natural rights of individuals and of education itself, within an organizational context; and at the practical end are the immediate needs, values, and desires of all constituents. True democratic methods are designed to control the tendency of the latter to destroy the former, maintaining a

productive tension between the two. Flawed democratic methods based on relativistic consensualism cause the tension to be disturbed.

If one extrapolates Bloom's argument into educational organizations, improvement must be grounded on moral values, and employ democratic methods only to discern whether they are being maintained in the practical context. It is the administrators fundamental role to make changes in that context when such values are threatened (Holmes, 1987).

Value emphases and priorities determine administrative patterns and policies, but these expressions of value can differ from context to context. The determining factor as to whether the administrative actions are appropriate is whether or not the values that they are intended to serve are in fact promoted, rather than destroyed.

A third method by which management of meaning is created is through information brokering. Administrators function as brokers who exchange, highlight, and interpret information for others (McClintock, 1987). They often become spokespersons for the perceptions of others and they often try to find common ground that will facilitate action. This is a function of their unique role placing them at the confluence of all interest groups.

In education this responsibility entails developing, amongst all sorts of different interest groups, an accurate picture of the accomplishments and achievements of the education system. This challenge is complicated by the reality that schools serve many publics, and sometimes contradictory flows of images must be communicated to them in order to create an accurate picture (Sergiovanni, Burlingame, Coombs & Thurston, 1987). In addition, public impressions of education are usually piecemeal and inaccurate, based on:

...a multitude of small, individual experiences acquired personally through direct contact with schools, or with children, learned through the media, heard from one's neighbors, or gained perhaps from an acquaintance who walks in the school (Carol & Cunningham, quoted in Sergiovanni, Burlingame, Coombs & Thurston, 1987, p. 31).

Impressions are also a legacy of one's own school years. However, impressions are also changeable. The task facing the administrator is to create accurate impressions of schooling and related student achievement. As McClintock (1987) states:

...The essence of this process is to use, and govern the use of, language and other symbols, stimulate action from which information about reality will be produced, facilitate interactions among stakeholders in order to share perceptions, and conduct retrospective interpretations of events. The sum of these activities is the management of meaning which is a core responsibility of the administrator's information and communication responsibilities (p. 319).

Management of meaning also tends to minimize organizational fragmentation by: (1) phrasing programs goals and purpose in a manner acceptable to all constituencies; (2) providing suggested alternatives for action that are commensurate with the value positions of affected individuals or groups; and (3) monitoring the outcomes and process of decided upon improvement strategies.

Sergiovanni, Burlingame, Coombs & Thurston (1987) emphasize the importance of this administrative role. They claim that it is a vital task in the maintenance of public confidence:

...(Administrators) need to communicate a flow of images to important constituencies to maintain their legitimacy in the eyes of such constituencies. For example, the public expects schools to operate in a rational, logical manner, and legitimacy in the eyes of the public requires that such an image be communicated clearly and decisively....the management of public confidence requires deft handling of the flow of images communicated to various constituencies (Sergiovanni et al, p. 32).

Earlier it was stated that developing a communal vision, greater empowerment, and better information brokering systems in an educational organization are all forms of this task of "managing meaning". The challenge, is of course, to design and/or alter the leader behavior and administrative policies and procedures so as to emphasize values appropriate to the organization, and to create a new ideology that empowers and dignifies the purposes of education. Evaluation, and particularly formative evaluation, is a process in which these skills are required. A first step in both administration and evaluation is to develop a metaphoric image that reflects the qualities and values of the organization that are desired.

The Role of Metaphor in Administration

Morgan (1987), House (1983) and Eisner (1978) suggest that a method of ensuring that an organization remain true to its values is for it to be conceived of and constructed according to a metaphor embracing its ideal characteristics. House describes the power of metaphor in the following way:

...Until recently, the employment of metaphor was thought to be merely ornamental. Metaphor was used to make an expression more poetic or to emphasize a point rhetorically. However, novel experiences usually are structured in terms of more concrete ones, and cultural notions in terms of physical ones. Metaphor is essential to our most complicated thought processes and a vital intellectual tool that we use to understand the world (p. 6).

and:

...social problem solving is mediated by the stories people tell about troublesome situations. The framing of the social problem depends on the metaphors underlying the stories, and how the problems are framed is critical to the solutions that emerge....Therefore, the underlying metaphor gives shape and direction to the problem solution (p. 7).

Morgan (1987) in his book Images of Organization supports House's contention. He claims that:

...our theories and explanations of organizational life are based on metaphors that lead us to see and understand organizations in distinctive yet partial ways....the use of metaphor implies a way of thinking and a way of seeing that pervade how we understand our world generally (p. 12; emphasis in original).

He goes on to present eight different metaphors that can be used to characterize organizational structure. Of particular interest to the discussion here are the "organization as machine" and the "organization as brain".

The Educational Organization as Machine

Traditionally organizations, even school organizations, are imaged by most theorists as mechanistic systems (McPherson, Crowson & Pitner, 1987). Indeed, the traditional bureaucracy constructed according to the principles of specialization, hierarchical authority, rule-oriented procedures, rationality, formalization, and the separation of role responsibility from personal

responsibility (Etzioni, 1964), is epitomized in the assembly line approach in mass production factories: employees are expected to act as machines, and behave as if they were parts of machines.

This mechanistic model is the predominant image used to characterize processes of schooling. Eisner (1979) claims that:

The dominant image of schooling in America has been the factory and the dominant image of teaching and learning the assembly line. These images underestimate the complexities of teaching and neglect the difference between education and training (p. 262).

He goes on to argue, as did Morgan, that "The metaphors... of schooling...that we acquire have profound consequences for our educational values" (p. 261). He strongly suggests that the image we hold to characterize the processes and functions of an organization can describe the qualities found in that organization, but at the same time, they also begin to change the organization: in a sense, they become self-fulfilling prophecies. Eisner (1979) claims, for example, that the industrial model as applied to education has emphasized rationality and scientific methodology to the detriment of many other important qualities of learning. He goes on to argue that these two principles lead to an emphasis on measurement in traditional evaluation methods, and these have "profound effects on the content and form of schooling" (p. 267). Effects, which he adds, are quite negative: "the quality of the student's experience has been generally ignored or seriously neglected" (Eisner, 1979, p. 268).

The Educational Organization as Brain

What are the implications if one views, and attempts to design, the educational organization as a brain? What influence might this metaphor have on administrative procedures, considering that one might consider the administration in fact the "brain" of the educational organization?

This vision emphasizes the importance of information processing, learning, and intelligence, and provides "a frame of reference for understanding and assessing modern organizations in these terms" (Morgan, 1987, p. 14). Of particular interest to this paper are two main qualities the image of the brain suggests for organizations: (1) "Learning to learn", the

necessity for building in a capacity for quality learning in an ongoing way; and (2) designing this capacity as holographic learning--whereby each piece of the whole is infused with the capacity of learning, and the qualities of flexibility and innovation are espoused. These qualities seem particularly appropriate for "loosely-coupled" organizations such as schools and school districts. In addition, the whole concept of viewing an organization as a holographic entity is one of the fundamental cornerstones that underlies the subjectivist paradigm.

"Learning to learn" refers to the concept that a system should constantly assess, reassess, and adjust its processes almost intuitively so as to be instinctively responsive to new demands from the environment. It implies a flexibility and openness to change and creative thinking that traditional bureaucratic organizations are not noted for. Freedom to analyze existing conditions, judge their suitability, and suggest improvements are characteristics that must be deliberately fostered within an organizational context; it is consistent with the problem-solving role of administration suggested by Liethwood (1986).

The learning capacity of an organization refers to its ability to establish systems to review and analyze its effectiveness and then adjust operations to keep pace with changes in its environment (see Figure 2.4). The diagram in Figure 2.4 refers to the combination of both evaluation and leadership as

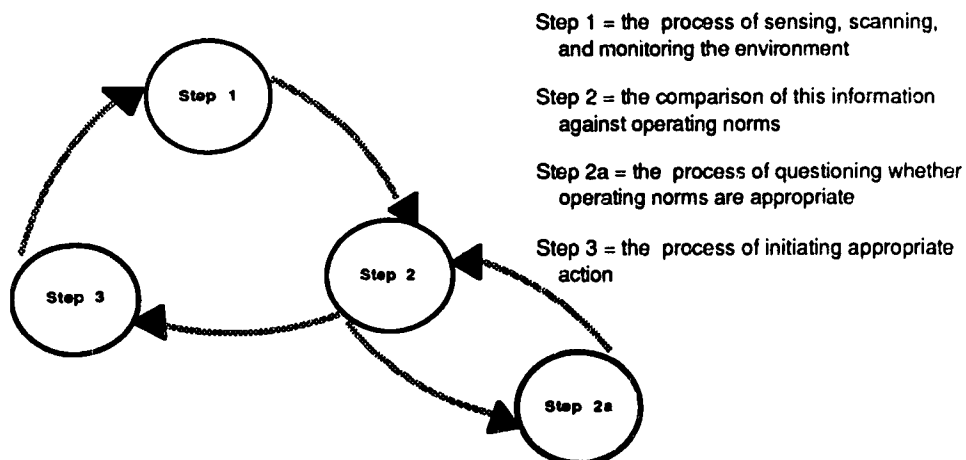


Figure 2.4: The "Double-loop" learning capacity of the brain (from Morgan, 1987, p. 88).

intuitively exercised by the brain. Steps 1, 2 and 2a are, in fact the stages associated with evaluation (Alkin, Daillak & White, 1979; Cronbach, 1982; Scriven, 1983a; Stake, 1982); step 3 is the manifestation of the leadership role of decision making (McPherson, Crowson & Pitner, 1987). Morgan (1987) argues that the traditional bureaucratic structure delimits the efficacy of these two important organizational functions, otherwise known in systems theory as maintaining "dynamic equilibrium" (McPherson, Crowson & Pitner, 1987).

In contrast, the brain accomplishes this two step learning process naturally. Organizations that develop openness, reason, and reflectivity; that encourage an approach to the analysis and solution of complex problems through the exploring of different viewpoints; and that devise means where intelligence and direction can emerge from ongoing processes, closely approximate the higher learning process of the brain. This makes it incumbent on such organizations to "...make interventions and create organizational structures and processes that help implement the above principles" (Morgan, 1987, p. 95).

In other words, effective educational administration will encourage and employ personal leadership and leader-substitute processes that develop a problem-solving capacity within that organization: the ability to intuitively adjust, monitor, and adapt the organization's processes instinctively when faced by new challenges either from the internal or external environments.

The brain learns both intuitively through naturally occurring experience, but also through artificially created learning experiences. Indeed, the whole existence of a school is predicated on the concept that learning can be enhanced by putting people through carefully constructed exercises that provide the qualities associated with effective learning. Hodges (1982) provides an example of a model of effective teaching constructed to maximize the potential for learning in the classroom. She capitalizes on the basic premise that teaching for learning, as constructed in the classroom, goes through four main stages (see Figure 2.5).

Stage 5 and Stage 1 of Figure 2.5, and the link between them, relates to the natural "double loop" learning process employed by the brain. The degree to which education systems employ effective evaluation methods and utilize those

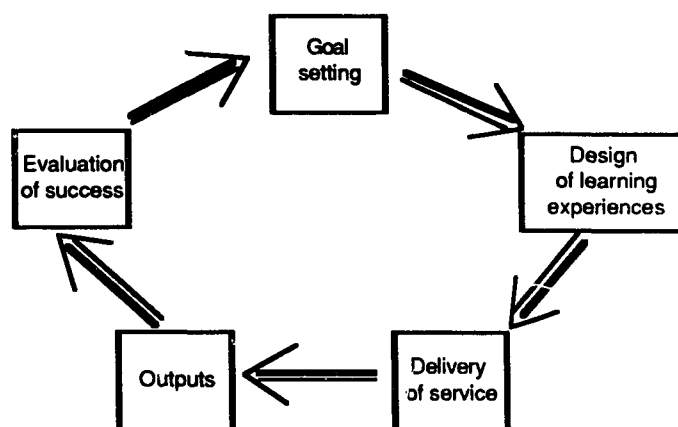


Figure 2.5 Artificially constructed learning loops, adapted from Hodges, 1982, p. 29.

methods for directing leadership action is to a great extent a measure of the degree to which their problem-solving capacity is enhanced. Effective organizations must establish policies and procedures for evaluation that ensconce these stages of artificial learning into day-to-day practice, just as they are fundamental to classroom learning. It is the function of administration--that part of the organization charged with the responsibility for leadership and service--to construct a strong problem-solving capacity within all components and members of that organization.

The greater the consistency of process and procedure from the smallest organizational unit to the largest, the greater the likelihood that the organization will act in active-adaptive fashion, maintaining a symbiosis with its environment. Thus organizations must have evaluation and leadership policies and procedures at all levels--built into habitual use--in order to be effective. These structures will act as leader-substitutes in enhancing administrative effectiveness.

An Educational Organization as a Holographic System.

Holographic systems are constructed so that the smallest unit of the whole has the capacity to assume the qualities desired of the whole itself (Morgan, 1987; Lincoln & Guba, 1985). Just as a piece of a holographic picture print can create the whole picture by shining light through it (and a three dimensional picture at that) so is it desirable that each section--piece--of an organization can create, upon demand, the problem-solving, or active learning capacity

described earlier. There are four main things that have to be done to create holographic organization:

1. Get the whole into parts.
2. Create connectivity and redundancy.
3. Create simultaneous specialization and generalization.
4. Create a capacity to self-organize (Morgan, 1987, p. 97-8).

Morgan suggests that there are four means that can be employed to build these characteristics into an organization. They are (1) redundancy of functions; (2) learning to learn (discussed earlier); (3) requisite variety; and (4) minimum critical specification. They are summarized in Figure 2.6.

1. Redundancy of functions. In mechanistic systems each part has a specific function and each new function requires an additional part. In a holographic system extra functions are added to each of the operating parts--

Holographic design rests on implementing four interrelated principles:

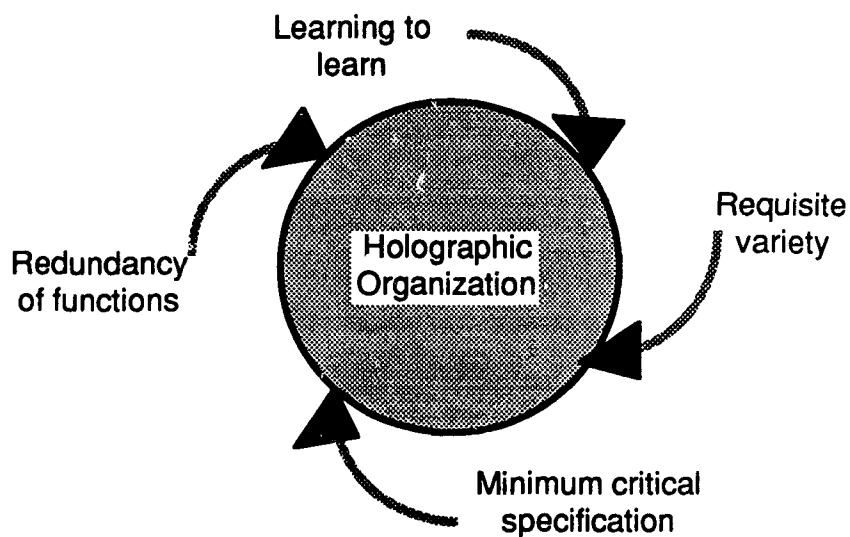


Figure 2.6: Principles of holographic design (Morgan, 1987 p. 99).

so that each part is able to engage in a variety of functions. Capacities relevant for the functioning of the whole are built into the parts. This implies that each organizational member has multiple skills and abilities; that in addition to their specialized function (i.e., teacher, or trustee) they are able to participate in any and all other activities of the organization when required.

If that activity is evaluation--then the responsibility for assuming that role is undertaken by all constituents involved in the educational process. This concept is encapsulated in the principle of "stakeholder involvement" in evaluation (Ayers, 1987); and participative decision making in leadership theory (McPherson Crowson & Pitner, 1987).

2. Requisite variety. It is, however, implausible to assume that everyone is skilled in all required areas. This is where the principle of requisite variety provides guidance. The basis of this concept is that particular roles within organizations face differing demands according to their relationship both to one another and the environment. Thus individuals must only possess enough--a requisite amount--of the needed skills to meet the demands their particular position places on that skill. This can be "built in" to the system by hiring individuals with a great multiplicity of abilities, or by structuring work groups so that a variety of skills are found in autonomous working units. In building this principle into evaluation, one would structure evaluation processes so that all constituents were involved, according to the idea of "stakeholder involvement"; and create tasks for individuals and design work teams so that their involvement was commensurate with their skill and perspective (Ayers, 1987). This component, combined with the principle mentioned earlier, redundancy of functions, is what is sometimes termed transformational leadership: empowering people to assume responsibility commensurate with their role in the organization.

3. Minimum critical specification. This principle reverses the bureaucratic principle that organizational structures need to be defined so that each hierarchical role has precisely and uniquely defined responsibilities. Organizations so constructed lose the capacity for self-organization. Minimum critical specification is consistent with the leadership concept of "purposing" (Barnard, 1938; Stout, 1987) by which guidelines structuring employee participation are loose enough to encourage self-initiation and ingenuity on behalf of the organizational member, but not so loose as to divert the thrust of such behavior away from the general goals of the organization. This principle means that hierarchical leaders strive to delegate "traditional" leadership functions to other members of the organization. Evaluation, instead of being conducted by the hierarchical administrator, should be structured to involve other members of the organization.

To this point in the chapter the literature in the area of background to the IPS, administrative philosophy, organizations, leadership and administration, and effective administration--consistent with the principles of the emerging paradigm--have been discussed. It suggests that the leader-substitute construct is a concept that expresses these principles. It also suggests that an appropriate metaphor for the educational organization is that of a learning organization, and that administration within that organization be considered to be the "brain" entrusted with the responsibility of building in the capacity of problem-solving. The administrator is responsible for establishing processes that encourage morality, vision-sharing, empowerment and effective management of meaning through emphasis on values and values resolution. A key process in achieving these goals is formative evaluation. Formative evaluation, constructed according to the principles of a holographic organization, is seen to be an effective mechanism for influencing organization members to become better problem-solvers.

Entrusted with the responsibility to initiate such influence, the administrator is encouraged to design a conceptual system in which tasks, structures, and characteristics of subordinates must be utilized to augment the administrator's own influence. Of particular interest to this dissertation is, of course, applying these principles to evaluation of school district operations. In order to accomplish this, it would be appropriate at this time to examine more closely the whole concept of evaluation, and what is meant by the terms "holistic", and "naturalistic".

Part IV: The Evaluation Function of the Information Profile System

To be an efficient and practical evaluation tool at the school district level the IPS must be designed to achieve the goals of effective administration: vision-setting, empowerment, and management of meaning. However, there are many possible evaluation strategies, each with its own philosophical base and purpose. In the remainder of this chapter a variety of evaluation approaches will be examined, with particular emphasis on holistic and naturalistic evaluation. Their methodologies suggest innovative ways to conduct school district evaluations, consistent with the principles discussed earlier. In addition, the naturalistic methodology is employed as the basic

research methodology for the overall study. A second part of the chapter will deal with limitations to evaluation effectiveness.

What is Evaluation?

Evaluation is a term that has many definitions and understandings. The Webster's Dictionary (1977) describes it as a 'determination of, or fixing the value of an object'; 'or the determination of the significance or worth of an object by careful appraisal and study'.

The Difference Between Evaluation and Measurement

Green (1970) describes evaluation as "the process of subjective appraisal with specific purposes or aims in mind" (p. 4). In most cases this appraisal is based on information or data which has been collected through measurement, which he describes as "the application of an instrument or instruments to collect data for some specific purpose" (ibid, p. 4). Evaluation is the more sophisticated process of the two; unless intuitively done, it requires some form of measurement. Evaluation always employs personal judgment and therefore is subjective in nature; on the other hand measurement tends to be as objective as the accuracy of the instrument and the skill of the user will allow. Green also states that evaluation "may lead to conclusions and applications which ultimately change that which is evaluated" (p. 13).

The difference between measurement and evaluation is an important one. Measurement is an objective, standard driven task; it is concerned with how much of a quality or characteristic an individual or object possesses. In this sense it appears to be nonevaluative, in that when something is measured, an indication is sought only as to how much of the characteristic is present. The measurement process puts no value on the amount. It is only through the process of evaluation that value "is added by the person interpreting the score" (Brown, 1971; p. 9). Measurement is therefore very selective and theory-bound; and it is only the beginning stage of evaluation, as determinations of value (merit or worth) are its essence.

In the book Standards for Evaluations of Educational Programs, Projects and Materials, developed by the Joint Committee on Standards for Education Evaluation (1981) this distinction is maintained. Evaluation is defined as "the

systematic investigation of the worth or merit of an object, i.e., a program, project, or instructional material" (p. 152). Measurement, in contrast, is described as the gathering of "quantitative information consisting of facts and claims that are represented as numbers" (p. 127). Adams (1964) restates this dichotomy in thinking when she states that "evaluation goes beyond measurement in that value judgements are involved" (p. 5).

Measurement implies knowing exactly what one is looking for, an ability to quantify it, and a comparison between a known amount (an accepted standard) and an observed reality. At this point measurement stops, and evaluation begins. Evaluation includes the measurement stage, but then adds a determination of the worth or value of the performance. It answers the question, "How well did we do" in terms of a person's satisfaction, as opposed to the discrepancy in performance itself. Measurement is clearly the process that applies when one has a clear definition of the objectives of an educational system, (i.e., specific test questions) and employs instruments to determine whether they have been achieved or not. Evaluation goes one step further. Using some internal or external ideal performance standard, one then determines the worth or merit of the achieved score. In so doing evaluation becomes a bridge between philosophy and practice, for measurement (the first stage of the evaluation process) captures performance in light of the world of practice, and then the data so collected is then compared to the ideal as articulated at the philosophy end of the administrative continuum.

Evaluation Research as an Administrative Tool.

Cooley and Lohnes (1976) add to the debate on evaluation by identifying evaluation as a tool for administrative purposes. In stating that educational procedures are never completely, finally evaluated, they imply that the fundamental purpose of evaluation in education is formative: dedicated to improvement of the educational program. They state:

An evaluation is a process by which relevant data are collected and transformed into information for decision making. Evaluation is defined as a process rather than a product... Evaluation is successful insofar as the information it generates becomes part of the decision making processes in education (p. 3).

Although they have a bias toward evaluation as conducted by official "evaluative researchers", as opposed to by administrators themselves, they make a number of points that are valuable to repeat here. The first is that "the omnipresent client of evaluators is society, and all evaluations which are intended to affect education policy should be reported to the public in suitable fashion and suitable media" (p. 4). A second is that evaluators must learn from their clients as well as informing them. Third, they focus on evaluation of student outcomes as the source of information from which evaluations of education should be determined. A fourth point is that evaluation provides a mechanism whereby assumed means-end relationships (i.e., procedures and policies derived to accomplish educational purposes) can be brought into the open and their validity examined.

Stufflebeam and Webster (1988) lend credence to this point of view when they define evaluation as follows:

Evaluation is the process of delineating, obtaining, and applying descriptive and judgmental information about the worth or merit of some object's goals, plans, operations, and results in order to guide decision making, maintain accountability, and/or foster understandings (p. 571).

In doing so they emphasize that evaluation employs judgement (i.e., determinations of worth and merit); it involves the gathering of data on a large number of variables; it has many purposes; and it is an interactive process involving complex communication and technical skills.

Values and Evaluation.

Stufflebeam & Webster (1988) also make the point that "evaluation is an assessment of value" (p. 571). Cooley and Lohnes discuss this topic at length, suggesting a "theory of valuation" (p. 9). They state that attempts "to arrive at information useful to others in considering the value of educational programs require some framework in which to think about the valuing process" (p. 9). They look closely at John Dewey's work to establish such a framework.

Cooley and Lohnes go on to make the statement that "it is high time for evaluators to recognize and assert that value statements can be analyzed into a set of propositions subjectable to empirical investigation" (p. 10). They also

claim that "failure to perform such analyses in evaluation studies is inexcusable" (p. 10).

Cooley and Lohnes' argument is as follows. First, the values driving educational practice must be revealed. Second, they make the assumption that in doing so some controversy over values will result; as a consequence they suggest that rational empirical inquiry should be utilized to settle such disputes; the "alternative...is the settling of value disputes by power alone" (p. 10). A third postulate of their argument is that values statements are propositions about matters of fact, and therefore drive both the means and the ends of educational endeavour. They state that the "primary task of the evaluator is to do research which generates information about the validity of propositions relating educational means and ends" (p. 11).

Evaluation, then, is a value-infused process. First, values determine which evidence is gathered so as to make determinations as how to describe the current state of educational programs. Second, one compares the state of educational programs in light of their ability to achieve the values of society that are "prized, held dear, viewed as invaluable ends" (Cooley & Lohnes, 1976, p. 14). In doing so one is then led into a third set of value judgements: making determinations as to what means can be used to alleviate undesirable conditions so found.

Cooley & Lohnes (1976) go even further: they make the claim that evaluations "are examples of moral behaviors" because they "seek to provide information germane to decisions about what is best for children, or for students of any age" (p. 16). In doing so they introduce two key elements of an effective evaluation exercise: first, an emphasis on students, both in light of the means and ends desired; and second, the use of a values framework that reflects society's moral principles, as these are the "invaluable ends" of educational action that must be converted into means and ends according to community need.

The Epistemological Debate in Evaluation

The epistemological debate in evaluation is a function of the ongoing philosophical debate with regards to the nature of reality, and truth. Three issues are of particular interest to this dissertation: the controversies surrounding independent reality versus dependent reality, qualitative versus quantitative inquiry, and positivist versus phenomenological research design.

Independent versus dependent reality

Two questions in particular need to be addressed in order to deal with this debate. They are: (1) What is the relationship between facts and values? and (2) What is the relationship between investigator and investigated? (Smith, 1983). The substance of the two questions as they relate to the issue of the educational evaluation is this: Can education be described as a valueless, free-standing entity, existing totally independent of the knower, or is it mind-dependent, and grounded in the individual observer's interests, values, and situation (see Fig. 2.7)?

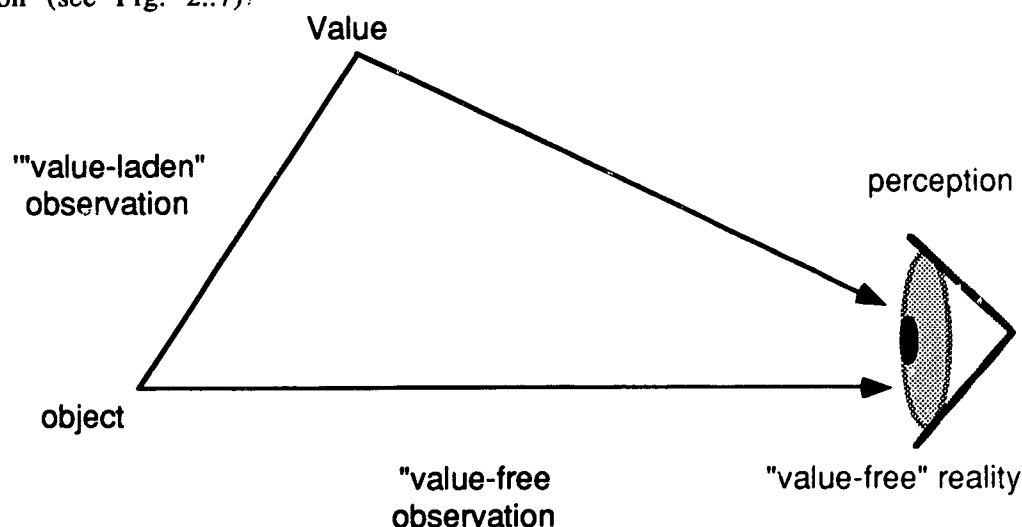


Figure 2.7: "Value-free" reality versus "value-laden" perceptions.

If one accepts the value-free position, then judgements can be performed, during evaluation, by comparing the educational system against some independent standard that exists on its own. Judgement then becomes objective, and substantial agreement among evaluators is the goal. If educational truths are mind-dependent, however, then judgements become a

function of comparing individual perceptions against a personally value-determined, ideal standard. Unanimity, in the latter case, is very unlikely. Implicit in the value-free argument is an acceptance of quantitative, objective truth, which is the basis of traditional scientific inquiry. It is founded on the positivist view that knowledge of an entity--say the educational process--is gained by breaking the complex into its component parts, and continuing in this division until its true nature is revealed. Understanding of the parts can lead to an understanding of the whole.

The value-laden approach--the qualitative, subjective viewpoint--presupposes, on the other hand, that knowledge of a complex form can best be attained by seeing it from a variety of viewpoints. In piecemealing it down, the essence of the object may well be lost, and its true nature elude the observer. If educational systems are realities, experts trained in scientific method and observation will be the best evaluators, as worth will become measurable and quantitative; if educational knowledge is perceptions, best determined by careful observation, anyone in the position of observing the evaluand can evaluate, using their own internal standard of worth, intrinsically or extrinsically determined.

There are many qualitative evaluators who do not accept the positivist view that even science--supposedly the most objective of all inquiry--is value free. Michael Scriven (1983a), for example, states that "science is about as value-free as buying hogs or playing chess" (p. 79). He argues that the scientist must judge between sound and unsound observations, valid and invalid generalizations; distinctions that demand value judgements that cannot be reduced to mere observations.

Scriven's value-laden characterization of science, when combined with the arguments of Lincoln and Guba (1985), are convincing. The latter state that positivism rests upon at least five assumptions that are difficult to maintain. They demonstrate cogently that the positivist belief in single, tangible reality, separation of knower from known, the temporal and contextual independence of observations (i.e., what is true at one time and place may also be true at another time and place), linear causality, and value-free inquiry are impossible to sustain in light of modern scientific evidence itself! They quote Hesse, who states that "all attempts to find demarcating criteria, that is, necessary and sufficient conditions for a belief system to be science have

failed" (Lincoln & Guba, p. 25). Ultimately one cannot define conceptual models in dissertations, and deal with topics such as evaluation without coming to a decision as to whether one adopts the concept of value-free inquiry or one does not. In the arena of evaluation, which is so substantively dependent on value judgements, this decision is not so difficult to make: however, an assumption throughout the remainder of the paper is that even the most basic of supposed "facts" in evaluation can only be considered, at best, subjective perceptions.

A further point. Adoption of the solipsist viewpoint does not mean that that one must retreat from the world and not know it. Such a viewpoint can equally argue for empiricism as against it (Barber, 1988). Indeed, empiricism can be construed as affecting "to know the world by reducing it to a measure of the self even narrower than a priori ideas, namely sensations or precepts" (Barber, 1988, p. 37). Independent reality or dependent reality notwithstanding, the only way one comes to know the world is through engagement with it. Even though certitude may not be possible, empiricism is a mechanism by which one can at least be led closer to an approximation of the truth, and a process whereby individuals seek to reduce perceptual differences through comparing observations and perspectives. Inductive empiricism is equally as important to non-science as science, for it leads us to a clearer picture with the non-realities we are dealing with (Barber, 1988). In a value-laden inquiry such as evaluation, this simply leads to the support of a variety of mechanisms for data collection, data interpretation, and construction of meaning that is based on a paradigm of inquiry constructed from the "non-independent reality" ontology, namely, naturalistic inquiry.

Schwartz and Ogilvie, quoted in Lincoln and Guba, (1985) explain how values do not necessarily mitigate against objectivity. They state:

There may, indeed, be an ultimate reality. However, every time we try to discover what it is, our efforts will be partial. Thus we see a shift from the "absolute" truth discovered by the "right" method toward a plurality of kinds of knowledge explored by a multiplicity of methods" (pp. 56-7)

Objectivity is gained by allowing many individuals to contribute to the characterization and expression of the the educational system. Truth is approximated by overlapping multiperspectives. This process brings sameness

and similarity to the fore, while at the same time dignifying individual viewpoints. These principles, in combination with a "value-laden" philosophy, and the principles mentioned earlier--constituent empowerment and participation--are the underlying concepts upon which the qualitative methodology and naturalistic inquiry are based.

Qualitative Strategy and Naturalistic Inquiry

Qualitative strategy is an approach that has its ultimate expression in a methodology called naturalistic inquiry. Investigations in evaluation can range from the fully reproducible, fully controlled artificial study (i.e., an experimental design, with its attendant emphasis on validity and reliability) "to the essentially opportunistic naturalistic study" (Cronbach, 1982, p. 43). The latter is more suited to evaluands of complexity, in which human interactions and subjectivity is necessary for an understanding of the entity being studied.

There are various levels of decision making in an evaluation study, and the choice of the method of data collection, either quantitative or qualitative, becomes the choice of strategy. Patton (1980) defines evaluation strategy as:

...a plan of action. A strategy provides basic direction. It permits seemingly isolated activities to fit together; as it moves separate efforts toward a common, integrated purpose. An evaluation research strategy, then, provides basic direction for the evaluator; it provides guidance in selecting particular techniques or methodological practices for specific settings (p. 40-41).

Patton (1980) describes the roots of the qualitative strategy. It employs an inductive versus a deductive approach, whereby one is not constrained by predetermined hypotheses, but allows understanding to develop as patterns and relationships reveal themselves. Research is an event in which the researcher makes no attempt to manipulate the research setting, but allows the phenomenon being investigated to evolve in its naturally occurring context.

Naturalistic Inquiry

The naturalistic research strategy is applied in the field where the program is dynamic, changing, and not subject to the tidy control of experimental design. As Patton (1980) states:

The evaluator makes no attempt to manipulate, control, or eliminate situational variables or program developments, but accepts the complexity of a changing program reality. The data of the evaluation include outcomes, changes in treatments, and patterns of action, reaction, and interaction (p. 42-3).

In naturalistic inquiry the researcher must get close to the program rather than remain aloof from it. Methods such as interviews, questionnaires, and observations are utilized to discover what is happening, and then verify what has been discovered. This means moving back and forth between induction and deduction, between experience and reflection on experience, and between lesser and greater degrees of naturalistic inquiry (Patton, 1980). This method is diametrically opposed to the traditional deductive, experimental, aloof and objective quantitative approach traditionally associated with scientific methodology. As such, it is suitable for some evaluation models but not others; one approach it is particularly suited for is naturalistic evaluation.

The underlying philosophy of naturalistic evaluation is based on the anti-positivist arguments of Lincoln and Guba (1985). A first premise is that there are multiple, intangible realities which can be studied only holistically. A second point is that the inquirer and the object of inquiry interact to influence one another; this is particularly so when the object being evaluated is another human being. A third postulate is that the purpose of evaluative inquiry is to develop understanding; therefore, differences are as inherently interesting as similarities.

A cornerstone of the naturalistic viewpoint is the "value-laden" nature of evaluation. Naturalistic evaluation claims that inquiry is value-bound, in at least five ways:

1. Inquiries are influenced by inquirer values as expressed in the choice of a problem, and in the framing, bounding and focussing of that problem.
2. Inquiry is influenced by the choice of paradigm which guides the investigation into the problem.

3. Inquiry is influenced by the values which inhere in the context.

4. Inquiry is influenced by the choice of substantive theory utilized to guide the collection and analysis of data and in the interpretation of findings.

5. Inquiry is either value-resonant (reinforcing or congruent) or value-dissonant (conflicting). Problem, paradigm, theory, and context must exhibit congruence (value-resonance) in order to produce meaningful results (Lincoln & Guba, 1985, p. 38).

As such, naturalistic evaluation, in its conceptualization, wishes to consider the value positions of all the members of an educational organization, while at the same time assuming that some are going to conflict with others, and some are going to be better than others. Resonance of congruent values, and resolution of value conflicts, is the purpose of the naturalistic design.

The Naturalistic Approach to Evaluation.

Because of its value emphasis, naturalistic evaluation is a process suited to the evaluation of complex social entities such as education (House, 1983; Lincoln and Guba, 1985), because it hones in on social processes. Indeed, Alkin, Daillak and White (1979) state that:

This sensitivity of naturalistic research to social process is precisely what is called for in research on evaluation utilization. In addition, the focus of naturalistic research on the social actors' points of view and understandings will allow us to uncover the various subtle forms of utilization, forms which might escape the attention of the outsider but which are readily apparent to the inside actor. The participant in an evaluation situation, who spends all or a part of each working day dealing with an educational program or its evaluation, can help identify the long and short term effects of the evaluation, including effects which the "detached" researcher might not anticipate and might well overlook (p. 33-4).

The purpose of naturalistic evaluation is formative in nature, aiming at understanding, in a "hermeneutics" sense (Smith, 1983). Naturalistic evaluation stresses multiperspectives (Farley, 1987), participation of stakeholders (Ayers, 1987; Greene, 1988; Greene, 1987a), and value-pluralism (Greene, 1987b).

Naturalistic principles can be conflated to provide a picture of an educational evaluation process that is constructed in the following manner. The naturalist admits the role that values play in shaping the inquiry and

seeks out a variety of viewpoints (multiperspectives). These viewpoints are gained through the use of qualitative methods of inquiry: interview, observation, discussion, logs, content analysis of meeting minutes, etc.). The research is conducted by an individual evaluator (who Lincoln and Guba call the human instrument) who subjects him/herself constantly to reference checks (i.e., one goes back to the person and checks the adequacy of transcription of the person's viewpoint, and value-statements), peer debriefing, audit checks of data and process, and triangulation (multiple data sources). In this way the theory guiding the evaluation is constructed, and value dissonance minimized, since the subjects' constructions and the substantive theory are "both extracted from the data rather than laid upon them" (Guba & Lincoln, 1983, p. 323).

Inherent in naturalistic evaluation is the belief that the human instrument has great stores of tacit knowledge (i.e. the grand total of one's professional experience and knowledge) that qualifies him, in a value-laden environment, to be the ideal evaluator. This is because the human instrument can react to data immediately, explore atypical or idiosyncratic responses when they arise, clarify, adapt, respond and summarize on the spot, and grasp phenomena "all of a piece" (Lincoln & Guba, 1985).

Trustworthiness in Naturalistic Studies

Trustworthiness refers to the qualities of credibility (internal validity), transferability (external validity), dependability (reliability) and confirmability (objectivity) (Lincoln and Guba, 1985). These must be established for each case study to ensure the academic integrity of the study. Table 2.2 indicates the measures that should be taken in this regard.

The purpose of a naturalistic study is not to necessarily provide generalizations to an outside audience, but to gather insight into the unique happenings within the context being evaluated. Generalization to other contexts is only possible if in the eye of the outside reader some aspects of the naturalistic study make such generalizations obtain to their particular circumstance. In multiple case studies, further generalizability may be possible; as Alkin, Daillak and White (1979) state, "One indicator of generalizability may be the extent to which recurrent issues and patterns may be identified across the settings combined with the reader's critical assessment

of whether these issues and patterns are recognizable and common" (p. 35) within his domain.

Joanne Farley (1987) claims that conclusions in naturalistic evaluations are justifiable when they meet three types of acceptability criteria. First, empirical claims must be empirically accurate; that is events did transpire as they are described in the study. Second, the description of events must reflect an accurate reconstruction of not only events, but the meaningfulness of the events. Finally, she states that "the interpretive explanation and related conclusions must be normatively adequate (p. 350). In other words, the explanations given of various decision making processes conducted by participants in the evaluation must encapsulate the essence of the value choices inherent in those decisions.

Table 2.2: Procedures for ensuring trustworthiness (adapted from Lincoln & Guba, 1985, p. 328).

Criterion Area	Technique
<i>Credibility:</i>	<ol style="list-style-type: none"> 1. Field activities: <ol style="list-style-type: none"> a. prolonged engagement. b. persistent observation (meetings, workshops, visitations). c. triangulation (interviews, juries content analysis, logs, multiple investigators). 2. peer debriefing 3. negative case analysis (refining the evaluation conclusions until agreed upon by participants). 4. referential adequacy; and 5. member checks (checking data with participants, both during process and at end).
<i>Transferability</i>	6. "thick" case study description
<i>Dependability</i>	7. a dependability audit trail (putting data away for analysis, spot checks by debriefers).
<i>Confirmability</i>	8. a confirmability audit (terminal; similar to above).
<i>All of the above:</i>	9. the reflexive journal (kept by the researcher).

Stakeholder Participation in Naturalistic Evaluation.

Just as stakeholder participation is suggested for holistic evaluation, it is suggested for naturalistic evaluation. Greene (1987b) argues that:

...a successful practical naturalistic evaluation is contingent upon meaningful participation of significant user groups in the evaluation process and on the open expression and legitimation of the multiple value claims held by members of such groups (p. 329).

In addition, naturalistic evaluators must make conscientious attempts to synthesize data and implicit or explicit value claims that will lead to justifiable evaluative conclusions. The integrity and accuracy of this meaning must be maintained (Pearsol, 1987).

Evaluation Models and Qualitative Strategies

Evaluation strategies suggest methods of data collection and design. Evaluation models are derived from a particular epistemological base and are larger constructs based on combinations of different data collection methods and design. Evaluation purposes are the goals that evaluation in one or other of its different forms can accomplish. Naturalistic inquiry is more akin to a strategy, as opposed to a model, and as such can be employed within models that are receptive to its epistemological base.

Evaluation Models

House (1978) identifies eight evaluation models, depending on (1) epistemological background (i.e., subjectivist versus objectivist); (2) whether they are inductive (goal-free) or deductive (goal-based); and (3) their intended audiences, suggested methodologies, and anticipated outcomes. These are shown in Figure 2.8, and Table 2.3. House's work is helpful in that it provides a clean, conceptual picture of almost all the accepted evaluation models and their progenitors, but also stresses the difficulty of designing evaluations that

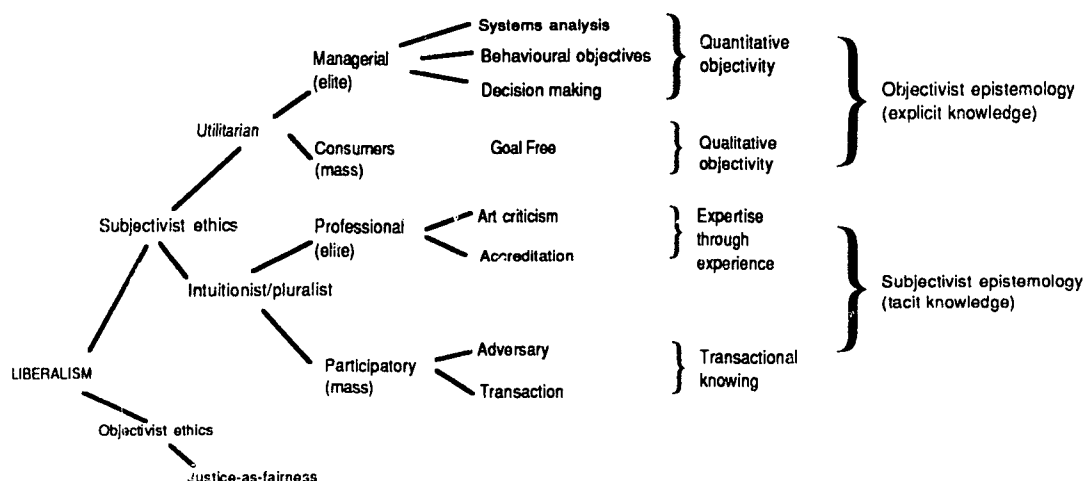


Figure 2.8: The epistemological base of evaluation models (From House, 1978, p. 50).

are not basically subjective in nature. He claims that all the major evaluation models are subjectivist in ethics; that is, the "ultimate criteria of what is good and right are individual feelings or apprehensions" (p. 50). This implies that judgements as to what is good or right are left to individuals, according to their own intuitive criteria, as opposed to being measured against an invariant outside standard, that somehow exists independent of people's value judgements. Even in sophisticated models such as the transaction model, in which pluralistic criteria are introduced by soliciting the judgments of various people in the program, data are weighted intuitively by the evaluator and the audience of the evaluation.

House even claims that systems analysis, behavioral objective and decision making evaluation models are also subjectivist. He argues that they are predicated on judgements based on utilitarian considerations; i.e., the choice of

Table 2.3: Evaluation models. Adapted from House, (1978), p. 49.

<i>Model</i>	<i>Proponents</i>	<i>Major Audiences</i>	<i>Assumptions Consensus on</i>	<i>Methodology</i>	<i>Outcome</i>	<i>Typical Questions</i>
Systems Analysis	Rivlin	Economists, managers	Goals; known cause & effect; quantified variables.	PPBS; linear programming; planned variation; cost benefit analysis.	Efficiency	Are the expected effects achieved? Can the effects be achieved more economically? What are the most efficient programs?
Behavioral Objectives	Tyler, Popham	Managers, psychologists	Pre-specified objectives; quantified outcome variables	Behavioral Objectives; achievement tests	Productivity; accountability	Are the students achieving the objectives? Is the teacher producing?
Decision Making	Stufflebeam, Alkin	Decision-makers, esp. administrators	General goals; criteria	Surveys, questionnaires, interviews; natural variation	Effectiveness; quality control.	Is the program effective? What parts are effective?
Goal Free	Scriven	Consumers	Consequences; criteria	Bias control; logical analysis; modus operandi	Consumer choice; social utility.	What are all the effects?
Art Criticism	Eisner, Kelly	Connoisseurs, Consumers	Critics, standards.	Critical review	Improved Standards	Would a critic approve this program?
Accreditation	North Central Association	Teachers, public	Criteria, panel, procedures	Review by panel; self-study	Professional acceptance	How would professionals rate this program?
Adversary	Owens, Levine, Wolf	Jury	Procedures and judges	Quasi-legal procedures	Resolution	What are the arguments for and against the program?
Transaction	Stake, Smith, MacDonald, Parlett-Hamilton	Client, Practitioners	Negotiations; activities	Case studies, interviews, observations	Understanding; diversity	What does the program look like to different people?

a measure, such as gross national product, or higher unemployment as a measure of success is still a measure of a positive choice on behalf of the evaluator. House argues that it is difficult to remove oneself from subjectivism in evaluation, regardless of the model or strategy one is using.

1. The CIPP model. Stufflebeam & Webster (1988) also review evaluation models, particularly within the context of evaluation as an administrative function in education. They discuss the CIPP model which consists of four levels of evaluation. The first is context evaluation, which assesses needs, problems, challenges or goals at each level of school district operations. The second is input evaluation, which seeks out alternatives and assess whether their application would accomplish the needs identified in stage one. The third level is process evaluation, which is a monitoring stage, in which implementation is guided and determinations made as to whether it was successful according to its design. The fourth stage is product evaluation--an attempt to examine the outcomes of a program and the extent to which they satisfy client needs.

Variations on this model are being employed in a number of school districts in the United States. Its strengths are its consistency from inputs to outputs and emphasis on context, which allows each district to deal with its own unique circumstances. Its challenges lie with the implicit assumption that one can define contexts, inputs, processes and products; and secondly, with the fact that it leads to idiosyncratic evaluation procedures that are only reflective of each individual district's situation--not allowing, therefore, for effective comparison district to district.

2. Formative and summative evaluation. Another model identified by Stufflebeam & Webster (1988) is the formative and summative model of Scriven, (1983b). Formative evaluation refers to evaluation directed at improvement, based on the assumption that education is an ongoing process. Summative evaluation is undertaken after a program's conclusion (i.e., to determine its merit or worth after the fact). Whereas summative evaluation can be applied to programs with finite duration, it is a difficult concept to apply to ongoing public education systems such as school districts and schools; and "evaluators who do an intelligent formative evaluation and retain the information are in a good position to summarize the information (and) provide program accountability reports" (p. 579) as well as give advice on decision-

making. Both the CIPP model and the formative evaluation model can be considered to be decisionistic models, one of the eight types defined by House (1978).

Evaluation Models Compatible with Qualitative Strategy

Patton (1980) uses House's models to analyze the applicability of qualitative methods such as observation, questionnaires, and interviews to evaluation design. He claims that the models of systems analysis and behavioral objectives are incompatible with qualitative methodology: "Qualitative methods derive from a different and conflicting paradigm" (p. 50). Art criticism, accreditation, and the adversary approach, being professional review models, are sometime compatible. And the transaction model, goal-free evaluation, and the decision making model are highly compatible with a comprehensive qualitative methods strategy. These models need to be considered in greater detail.

1. The transaction model. The transaction model of evaluation focuses on events occurring in and around the actual system or program being evaluated. It is based on perception and knowing as a transactional process, whereby meaning and creation of meaning is constantly being altered and reconstructed. It is predicated on three main principles (House, 1978):

- (1) the facts of perception are always presented through concrete individuals dealing with concrete situations;
- (2) perceiving by each person is always done from his own unique position, experiences, and needs, including the scientist-observer; and
- (3) each person creates his own psychological environment by attributing aspects of his experience to his environment.

In this model, a value-free reality is impossible, as knowing is a function of the interchange between individual and environment, gained through active participation in that environment.

2. The goal-free model. The goal-free model of evaluation is the product of Michael Scriven (1974). He warns about evaluations being so concerned with determining the success or failure of goals that they fail to take account of

unanticipated outcomes that might be as equally important to the enterprise. As he himself states:

...I became increasingly uneasy about the separation of goals and side-effects. After all, we weren't there to evaluate goals as such--that would be an important part of an evaluation of a proposal, but not...of a product. All that should be concerning us, surely, was determining exactly what effects this product had (or most likely had), and evaluating those, whether or not they were intended (p. 35).

Scriven claims that goal-free evaluation is best employed in the formative stage of evaluation, looking for the actual effects, rather than checking on alleged affects. In doing so he claims that evaluator bias is minimized, real effects can be disentangled from "anticipated" real effects, and rigidity of thinking minimized. He discounts skeptics who claim that adoption of this approach does not allow focused evaluation, leads to poor planning, is too broad to employ, and is a threat to many producers because it involves all people affected by the program in its determination. He does so by simply claiming that goal-free evaluation should in fact provide better information, in that it includes unanticipated effects, and should therefore lead to better planning, and improvement. In the long run, then, it should be less of a threat than an advantage to the administrator of an organization.

3. The decisionistic model. Floden and Weiner (1978) use the term decisionistic model to refer to evaluations designed for one specific purpose: to improve decision making. They state that conventional wisdom "concerning the relationship between evaluation and governmental processes emphasizes the impact of evaluation upon discreet decisions made by public managers" (p. 178). This model contains three fundamental assumptions: (1) that programs function to achieve clearly defined goals; (2) that evaluations serve to collect information relating to increasing knowledge about the effectiveness of programs in light of these goals; and (3) that this information will be used by decision makers to make discrete, identifiable decisions intended to improve programs, because of an inherent desire for improvement.

Both Floden and Weiner (1978) and Patton (1980) argue, however, that the decisionistic model need not be narrowly construed as to relate to only very specific goals, but can also be designed to gather information about a very broad future context within which decisions must be framed. As this

information can consist of aspirations, needs, or values of a variety of constituents in education, it is open to all kinds of methodology. The methods used depend on "what evaluative information is necessary to help make specified decisions" (Patton, 1980, p. 58). Qualitative methods therefore are appropriately employed in the use of this model.

Purposes for Educational Evaluation

In addition to the purpose of decision making for system improvement, evaluation can serve others functions equally well, particularly when applied to school district operations. Evaluation can be used to: (1) reduce conflict and overcome complacency (Floden & Weiner, 1978); (2) assist with accountability (Glasman, 1986; Stufflebeam & Webster, 1988); (3) enhance knowledge, i.e., be used as an educative tool (Cronbach, 1982); and (4) help with motivation and gain of public support (Glasman, 1986).

Conflict Reduction and Complacency Reduction

Evaluation can be viewed as a method for managing conflict so as to promote a "teamwork" approach to change. Evaluation is a signal that the program under study is subject to discussion, negotiation, and compromise. If conflicting groups agree to participate and contribute to an evaluation, then the likelihood of issues of conflict being replaced by a reaffirmation of a common purpose is enhanced. At the least, a shift in the debate from issues of disagreement to "evaluation strategies indicates that both sides of the debate have tacitly agreed to investigate at least one new program" (Floden and Weiner, 1983, p. 182).

Evaluation can also reduce complacency. Complacency refers to unhealthy self-satisfaction and unwillingness to change (Havelock, 1973). For change to take place complacency must be reduced, so improvement can begin. "The very act of participating in an evaluation may spur the consideration of new practices by practitioners...managers, [and] teachers" (Floden and Weiner, 1983, p. 182). Participation in evaluative efforts can prompt both the clarification of standard operating procedures and their revision, leading to greater clarity and more alternatives. This, in turn, leads to better decisions and an improved quality of life in the organization.

Evaluation for Accountability

A major theme of this paper is that political dynamics resulting from value conflicts over education have produced demands for accountability in the B.C. education system. Accountability is meant to ensure that those who design and operate educational programs carry out their duties responsibly, and provide learning experiences to students of a high quality. Coombs (cited in Calam and Fleming, 1988) describes accountability as a process by which the public retains control and authority over educational programs that they have delegated as a responsibility to educational professionals. The electorate must:

...impose on their agents an obligation to be accountable to them, i.e., to provide them with evidence for judging that the agents are acting responsibly in designing and operating educational programs. Accountability obtains to the extent that the agents accept and fulfill their obligations to provide such evidence (p. 69).

Stufflebeam and Webster (1988) explain how evaluation can be used to assist with accountability. Evaluation can give an accounting of what has been accomplished, why it was accomplished, and at what expense. They state that "a school system's staff could hardly address such issues to the satisfaction of...the public if they could not back their claims with a record of pertinent and credible information" (p. 571). Evaluation activities can provide valuable documentation that will serve the demands of accountability that schools and school districts must meet.

Glasman (1986) states that accountability demands arise from value conflicts in education. In British Columbia the multi-cultural nature of society, demographic and economic trends, and family structure alternations have created conflicting demands for education, and a multiplicity of values that must be sorted through if the system is to have a consistent direction (Calam & Fleming, 1988). As education is a process by which societal values are passed on to the younger generation, these demands have been transformed through the political system into educational policies. Calls for greater accountability (Royal Commission on Education, Summary Findings, 1988; p. 184) are a result of these tensions in society.

Glasman (1986) identifies three criteria for educational accountability. These are efficiency, equality, and quality. The efficiency criterion reflects a

need to relate outcomes, in terms of student achievement, to financial expenditures incurred in developing this achievement. The equality criterion reflects the deeply held premise in western society that members of disadvantaged populations must be provided with equal access to educational opportunity. The quality criterion arises from recent disaffection with the job schools are doing, and is grounded in the need for evidence of sound, systematic and objective proofs as to what the achievements produced by our expenditures are.

Efficiency, equality, and quality are, however, values in themselves: values reflective of society's expectations for education. Accountability directed at providing evidence relating solely to these will not, however, assuage the value conflict surrounding educational purposes. To this end values inherent in the conception of education must be championed and measured; as well as those values fundamental to a democratic society (Coombs, 1988). These values should become the "ideal" drivers of the educational evaluation model, and the source of comparison against which evidence indicating success should be gathered.

Evaluation for Understanding

Stufflebeam and Webster (1988) also make the point that evaluation can be used to promote understanding of educational programs. This goal is also appropriate to the circumstances existing today in education in British Columbia, because there is a perception amongst the educational community that a lack of awareness as to the true purposes and accomplishments of the system are responsible for the poor public perception of education, particularly amongst citizens who do not have children in school (McIver, 1985). Evaluation provides a tool to increase understanding, and is, indeed, itself an educative act. Cronbach (1982) makes this point succinctly:

...The evaluator...is an educator. His success is to be judged in his success in communication; that is, by what he leads others to understand and to believe. Payoff comes from the insight that the evaluator's work generates in others (p. 8).

Cronbach goes on to say that good evaluation cannot be defined to a singular audience in a political enterprise such as education. A variety of audiences must be considered both in determining what questions should be asked, as

well as how to report results. Stufflebeam and Webster (1988) agree, stating that "school system evaluators need to differentiate their audiences; determine their unique and common information needs; and design conduct, and report their findings accordingly" (p. 574).

As educational systems get more and more complex, in moving from the classroom to larger units of analysis such as school districts, "understanding" becomes a more and more important goal of evaluation. One version of understanding is hermeneutics (Smith, 1983). This term refers to an attempt to understand others through an interpretive study, by bringing one close to the experience and trying to re-create it in oneself. It refers to a knowledge of context or background that is necessary for the interpretation of events. Hermeneutics is a circular process in that every part of a whole requires the rest of the whole to make it sensible; yet the whole cannot be understood except in terms of its parts. Therefore understanding is accomplished by moving back and forth between the whole and the parts, and constantly directing oneself towards understanding of the dynamic relationships that define and delimit the object in question, in terms of their synergy with one another.

This form of understanding is reflected once again in the naturalistic form of inquiry. As one moves from the small, finite world of experimental science whereby all variables can be controlled and manipulated, to the complex domain of district educational systems, one is tempted to move to the naturalistic methodology. This is because it stresses: (1) not knowing, in advance, what is important to observe; (2) respect for each person's values from their own perspective; and (3) recognition that it is interpretations by program staff and clients that determine the outcomes (Cronbach, 1982).

Evaluation for Gain of Public Support

Earlier it was stated that evaluation can reduce complacency and dysfunctional conflict within an organization. This same principle can be applied to the voting public, if one defines the organization as consisting also of parents, community members, and the local electorate. In this sense evaluation attempts to gain public support for education.

Evaluation for public support is achieved by creating a new meaning and purpose for the education system amongst the electorate. By extending the

educative role of evaluation to the larger constituency, one creates a need for support and motivates this support by providing the public with more accurate, meaningful images of education: images that identify needs and strategies for achieving these needs that require public support for them to be successful. "Evaluation implies change" (Glasman, 1986) and therefore both a need for change and a direction for change can be created through the provision of evaluation information.

Sergiovanni, Burlingame, Coombs and Thurston (1986) stress that creation of new meaning and purpose is not an easy task. The challenge facing the administrator is to create accurate impressions of schooling, and related student achievement; but this challenge is complicated by the reality that schools serve many publics, and sometime contradictory flows of images must be communicated to them in order to create an accurate picture. This implies a variety of evaluation reports and communications strategies, all that take time and resources to develop.

Utilization-focused Evaluation

This approach to evaluation reflects a focus on utility for evaluation design--suggesting that epistemological consistency (i.e., adherence to a particular methodology or strategy for epistemological reasons) can limit effectiveness. The utilization-focused approach to evaluation (Patton, 1978) represents an attempt to move beyond theoretical models to the practice of evaluation. It is an attempt to apply problem-solving skills, (what Patton calls active-reactive-adaptive thinking) to evaluation challenges (Patton, 1980). Utilization-focused evaluation is an orientation to evaluation, not a model or methodological strategy.

Patton's Utilization Guidelines

The main caveats of this approach are described by Patton (1978) as follows:

There are only two fundamental requirements in this approach: everything else is a matter for negotiation, adaptation, selection, and matching. First, relevant decision makers and information users must be identified and organized--real, visible, specific, and caring human beings, not ephemeral, general, and abstract "audiences", organizations, or agencies. Second, evaluators must work actively, reactively, and adaptively with these identified decision makers and information users to make all other decisions about the evaluation--decisions about

research focus, design, methods, analysis, interpretation, and dissemination (p. 24).

The various steps suggested by Patton as the procedures in utilization-focused evaluation are summarized in Table 2.4.

Table 2.4: An outline of the Utilization-Focused Approach to Evaluation (adapted from Patton, 1978, pp. 284-88).

The Five Major Steps to Utilization-Focussed Evaluation

- I - Identification and organization of relevant decisionmakers and information users.
- II - The relevant evaluation questions are identified and focussed.
- III- Evaluation methods are selected that generate useful information for identified and organized decisionmakers and information seekers.
- IV - Decisionmakers and information users participate with evaluators in data analysis and data interpretation.
- V - Evaluators and decisionmakers negotiate and cooperate in dissemination efforts.

Patton also argues that developing evaluation strategies and models with a utilization focus depends on the use of any and all data that will help shed light on evaluation questions. The issue is not which ontological paradigm is the correct one, so much as which method will produce the information needed to achieve the result. He states that "the debate and competition between paradigms is being replaced by a new paradigm--a paradigm of choices" (p. 20). Utilization-focussed evaluation uses both qualitative and quantitative data when necessary to achieve its purposes.

Factors Affecting Utilization

Alkin, Dalliak and White (1979) identify eight factors that influence the potentiality that evaluation will be effectively utilized by an educational organization. The first factor is pre-existing evaluation bounds. Variables such as school and community conditions, the mandated bounds of an evaluation, and fiscal constraints all have a say in the success of the evaluation. A second factor is the orientation of the users. Their expectations for the evaluation, their pre-existing concerns about the program, and their

preferred forms of information affect the evaluation process. If evaluation is a learning process, then the mechanism used to impart knowledge will be a function of the learner's best learning style.

A third factor is the evaluator's approach. Both the model and the methodology employed alters the evaluation's utility. The role the evaluator chooses for himself also affects the results, as it defines also the role of stakeholder participation. So too do his choices within the bounds established for the evaluation, and his manner in carrying out the mandated tasks of the evaluation; i.e., rapport with other participants. This is particularly important for evaluators who seek to "facilitate and stimulate the use of information" (Alkin et al, p. 244).

A fourth factor is evaluator credibility. For evaluations to have impact, the trust and expertise of the evaluator must be established. Credibility in evaluating education, for example requires knowledge of the system, sensitivity to different audiences and their needs (i.e., teacher, community, administrator and trustee), and a demonstrated expertise in evaluation.

A fifth category were organizational factors. Relationships between the central office and schools; between management and teachers; and established policies and procedures, both formal and informal, affects the kind and availability of information for the evaluation. So too the expectations and aspirations for the evaluation held by the individuals commissioning it; this is reflected in their openness, their financial support, and their receptivity to the results.

Extraorganizational factors are the sixth category of factors. This refers to influences outside the school organization. Community influence in terms of interest and support, and the influence of other governmental agencies can alter the utilization of an evaluation.

A seventh set of factors are information content and reporting. The amount, kind, and method of presentation of evaluation information can influence its utility. Whether one uses a variety of different reports for various audiences or not may be a factor. Whether an ongoing or terminal dialogue is established can have an impact; "information dialogue results in a more profound and *shared* perspective on the uses of evaluation" (Alkin et al, p. 254; emphasis in original). Patton (1986) emphasizes this viewpoint, making the case that creative evaluation reporting can help capture the essence of an

educational program. The use of pictures, visual aids, etc. in a report can help create an empathetic response that captures the essence of the program being evaluated.

Administrative style is the final factor identified by Alkin, Dalliak and White (1979) as affecting evaluation utilization. Administrative and organizational skills and administrative initiative (i.e., a willingness to take charge and act) are two factors in this domain. In making the comment that "the 'right' combination of administrator skill and initiative can be the primary factor in keeping an evaluation targeted toward useful tasks and appears to have enhanced the possibility of utilization" (p. 256) Alkin et al lend credence to the concept that an evaluation model designed according to the concept of leader-substitutes and the dynamic relationship between them and the formal leader may enhance its potential for success.

Factors Limiting Evaluation Utility and Influence

A number of writers give reasons why evaluation has not achieved its theoretical promise in practical use. Indeed, this was the purpose of the Alkin et al study: to determine whether in fact evaluation made a difference, and what factors determined its ability to make a difference. Citing a number of authors who state, as do Worthen and Sanders (1973) that: "Evaluation is one of the most widely discussed but little used processes in today's educational systems", (p. 1) they summarize the literature until 1979 on this topic. They come to the conclusion that evaluation does, at times, come close to fulfilling its promise, and suggest a number of factors that delimit this potential.

A first point Alkin et al (1979) make is that sometimes evaluation doesn't appear to be utilized because those who are making these judgements employ too narrow a viewpoint as to what constitutes effective utilization. Too often utilization judgements are based simply on a judgement as to whether evaluation significantly influenced decision making. They suggest that the definition of utilization be expanded to include "influence on administrator perceptions of the evaluated program, which, in concert with other forces, may slowly change the course of program decision-making" (p. 25). This perspective on utilization emphasizes interpersonal interactions of people and situations in the evaluation process, and identifying the full range of potential influence on the future of the educational organization.

Alkin et al (1979) quote Weiss (1966) who summarizes a number of the practical reasons why evaluation does not necessarily have the influence on program decision making that it should. She cites methodological reasons, inadequate academic preparation, practitioner suspicion and resistance, problems with data access, design difficulty, inadequate time for follow-up, inadequacies of money and staffing, control on publications, and political machinations by senior administrators as possible constraints on evaluation usage.

Floden and Weiner (1979) agree with Weiss, suggesting that the problem is that evaluation has tied itself primarily to being a process of information gathering for the purpose of decision making, and has missed some of its potential functions, such as complacency reduction and conflict resolution. Stufflebeam & Webster (1980) claim that most evaluations are not, in fact, evaluations. They are pseudoevaluations; they have not been designed to be consistent with proper evaluation theory.

Scriven (1983b) indicates that the problem is attributable to overconcentration on goal achievement, leading one to ignore the "unanticipated" products of educational endeavor (the basis for the goal-free model discussed earlier). This leads to an incomplete understanding of the holistic nature of an educational program. It also leads to an overemphasis on process rather than outcomes, meaning that the true nature of the program's accomplishments (or lack of) is not understood. He also suggests that evaluations are too often designed to meet the needs of management, and are not consumer-oriented. He coined the term "valuephobia"--fear of evaluation--to describe the antipathy to evaluation that often limits the openness and honesty of participants (Scriven, 1983b; p. 230).

Politics are a final factor that contributes to the fact that evaluations often do not realize their intended goals. Stake (1983), for example, talks about cooption, caused when "the rewards to an evaluator for producing a favorable evaluation report often greatly outweigh the rewards for producing an unfavorable report" (p. 289). Another form of political influence reflects political necessity. For educational organizations, satisfaction ordinarily disappears only when external pressure cause substantial discomfort within the organization. Educators are content to maintain business as usual until public pressure focuses on problems in its domain. Then evaluation becomes a

suggested mechanism upon which to base improvement. However, the public is extremely fickle in its concerns. Once public excitement passes, the stimulus to improve services fades as well. Evaluations conducted in times of great controversy can be of little interest when completed (Floden & Weiner, 1983).

Stufflebeam and Welch (1986) conducted a meta-analysis of over 150 evaluations carried out in the past twenty years in the United States. The purpose of their study was to "provide information that could be used to assess and improve evaluation studies and evaluation systems" (p. 152). Their study is best termed a meta-evaluation (Nilsson and Hogben, 1983); an evaluation of evaluations. Meta-evaluations concern themselves with analysis of both the process and results of evaluations; in so doing they make judgements of an already highly judgmental process. Ultimately these judgements relate to the issue of utilization and influence, or overall evaluation effectiveness.

Stufflebeam and Welch (1986) found that there was very little conceptual pattern amongst the evaluations they studied. Most evaluations are "idiosyncratic and of local interest and not presented in a form for many" (p. 166). No report is given, for example of the use of a district-wide holistic evaluation model (i.e., one that encompasses all programs and procedures).

A major factor determining an evaluation's success is not so much the integrity of its design but its utility. "Regardless of the soundness of the methodology used, if the findings are not relevant to the problems of practitioners and used by practitioners, then the study may be judged to be a failure..." (Stufflebeam & Welch, 1986, p. 162). A major factor determining an evaluation project's utility is its efficacy in the eyes of the political masters who have commissioned it.

Other factors influencing utility are the commitment of the chief administrators, the involvement of the audience for the evaluation, the use of multiple measures, the use of a flexible and responsive approach to study, simplicity of reports, and externally mandated requirements for the evaluation (Stufflebeam & Welch, 1986, p. 162).

For any emerging evaluation system, then, two kinds of problems stand in the way of realizing its maximum return. The first relates to concerns that can be termed, "methodological concerns". These relate to the design of the evaluation structure: whether its scope is wide, to go beyond pure formative

evaluation and address unintended as well as intended outcomes, and issues such as conflict resolution and complacency reduction; whether it is outcome or process based; whether it is client or management centred; whether it is constructed to address immediate political whims or longstanding public concern.

The second set of factors affecting evaluation use and influence are political factors. Overcoming the limitations raised by these concerns is the leadership role inherent in evaluation. Clearly a major issue that will impact on the the ability of the Information Profile System to accomplish its evaluative tasks is the leadership skill of the administrators given the responsibility to carry out the process. Can educational leaders escape "co-option", overcome "valuephobia", and accept the need to consider clients--not only in relation to themselves, but for all other participants in the process (i.e., teachers, school board members and administrative officers, all of who who may be affected by an adverse, or unflattering evaluation)? Does evaluation have a role beyond being a response to issues of controversy, i.e., as a systematic, regular exercise in school district administration? These questions will be considered during the "action research" being done in this dissertation.

Holistic Evaluation as a "Utilization-focused" Strategy for School District Evaluation.

Holistic-inductive evaluation through naturalistic inquiry is a strategic approach employed to enhance evaluation utilization by school district administrators. It has been stated that holistic evaluation, employing wholly or partially naturalistic principles, could assist with the leadership functions of improvement, accountability, understanding, motivation and gain of employee and public support at the school district level of operations. An analysis of evaluation effectiveness has also suggested methods to enhance its use, as well pointing out limitations to its influence. It is appropriate at this time to explain how the holistic strategy is designed to overcome the constraints and maximize evaluation's usefulness for school district administrators.

What is Holistic Evaluation?

Holistic evaluation applies the principles of utilization-focused evaluation to large and complex objects such as district education systems. It stresses understanding the program as a whole, by seeking out subjective knowledge, such as constituent attitudes and values, as well objective knowledge of educational achievement such as test scores. Holistic evaluation employs an inductive approach, looking for actual effects of an educational program and making generalizations upon observed outcomes, rather than a deductive, goal-bound approach. The process stresses practical utility--by identifying the specific decision makers and information users, and constructing a process that involves all affected persons. In its truest sense, holistic evaluation educates all system constituents about other group priorities and interests for education. As Moran (1987) states:

This holistic model is designed to provide managers with information on how policy choices can affect the efficiency or effectiveness of their programs...The utility of this approach lies in the comprehensiveness of the research design, which provides policymakers with sets of data on how well a program is doing from bureaucratic, political, and societal perspectives (p. 613).

The Holistic Evaluation Process

Holistic evaluation, in theory, is a three stage process that is a composite of a variety of models (see Figure 2.9). The first stage has two steps. The first step is the articulation of a values framework around which data is to be collected. This framework is traditionally an educational philosophy, or mission statement produced by the agency governing the educational system under scrutiny: it is an educational ideal. The second is the collection of knowledge (data) about the object being evaluated. In holistic evaluation of large education systems, this knowledge is a composite of facts, attitudes, opinions, values and observations possessed by many people within the educational arena. This knowledge--quantitatively or qualitatively determined, through a variety of techniques--can be considered empirical in that it is the observable outcomes of the operationalization of an education

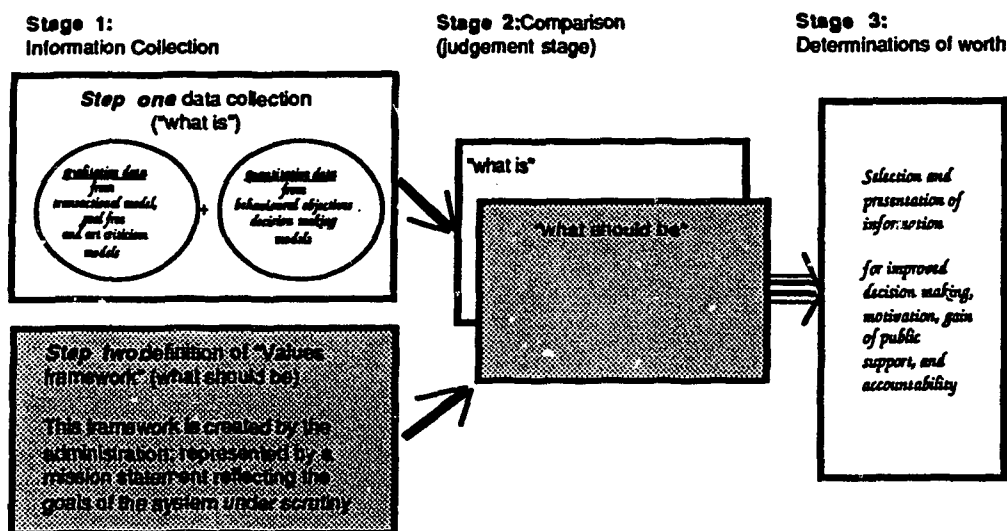


Figure 2.9: The three stages of holistic evaluation

system. Such knowledge serves as the data upon which the next stage of evaluation will be constructed. As Patton (1980) states:

...the holistic approach to research design is open to gathering data on any number of aspects of the setting under study in order to put together a complete picture of the social dynamic of a particular situation or program (p. 40).

The second stage of evaluation is to judge whether the information one has about the education system reflects a desirable state of affairs or not. In order to make such a judgement, one must compare what is (ie., the knowledge one possesses about the educational system) against the already articulated ideal standard. Comparisons against an ideal standard reveals discrepancies between what is and what should be. Evaluation then enters a third stage, also a stage exercising the quality of judgement. This is the making of determinations of worth--i.e., deciding which discrepancies are of greatest importance, and which values or issues need to receive priority attention. In keeping with the emphasis on the importance for administrators to translate the philosophy and values of an organization into practice, this implies the focus of holistic evaluation in a school district should be on determining the strength and weakness--the viability--of the district mission statement.

Wholey (1983) suggests an administrative approach called "results-oriented management" that captures an important focus of the holistic method but that

does not provide the philosophical emphasis required for effective administration. He states that "results oriented management" is a product of public demand for improved social service administration, particularly in light of spiraling tax demands. It is an attempt to utilize a number of procedures, procedures involving evaluation, to demonstrate to the public the accomplishments and achievements of the organization--and to provide a mechanism by which the organization can, through monitoring these outcomes, improve overall performance.

A strength in Wholey's (1983) approach is his emphasis on outcomes. He describes the process as:

- (1) agreement on a set of program outcome objectives and outcome indicators in terms of which the program will be assessed and managed;
- (2) development of systems for assessing program performance in terms of those outcome objectives;
- (3) use of program outcome information to achieve improved program performance; and
- (4) communication of program performance and results to policy levels and to the public (p. 7-8).

Defining management as distinct from administration, Wholey (1983) suggests this process as exclusively related to the practice end of the administrative continuum. As such the outcome indicators are a function of the behavioral objectives, or specific goals of the organization. In other words, outcome performance is used in this process of management to evaluate the suitability of the specific goals established for the organization; as such they are deductively sought out, and goal-based, versus goal-free. In addition, feedback to the ultimate philosophy of the organization does not necessarily take place. Wholey's model appears to be a reductionist model, whereby the whole is considered to be the sum of its constituent parts, and "mechanistic" in operation. On the other hand, holistic evaluation implies taking information about key components and reconstructing a clearer picture of the whole--which may be greater than the sum of its parts. Although the basic precepts of Wholey's theory relate closely to the IPS model, these differences conflict

with the basic principles describing effective administration and holistic evaluation as found in this paper.

Moran (1987) describes holistic evaluation as a process that "integrates program evaluation into the decision making of a public agency" (p. 613). The goal, he states, is to derive usable and comprehensible information about the achievements of a program through a variety of research methods that will relate well to the managerial issues the administrator faces. This paper suggests that a similar approach can be used to sort out the philosophical issues the school district administrator faces. Holistic evaluation can be constructed to help sort out decisions of philosophy, purpose, and value priority by gathering information about the achievements of the district's mission statement. In addition, it is suggested that it can be used to achieve the additional goals of improved accountability and the maintenance and gain of public support for education. A number of points relating the holistic evaluation concept to the literature review and to the existing IPS model need to be made at this point.

1. Values, administration, and holistic evaluation. A first point is to remember that holistic evaluation, like evaluation of any kind, is a value-laden process. Values, or concepts of the desirable, are the fundament of a district's "ideal" mission statement. They are the source of constituent opinions as to the success of educational programs. They become the source of criteria used to establish standards of acceptable performance. It is precisely this value-laden nature of evaluation that makes it such a valuable tool for administrators who are charged with the responsibility of maximizing rationality in an organization (Simon, 1971). Disparate values are a major impediment to rational decision making. Disparate values are why vision-setting, management of meaning, and empowerment are skills in high demand for today's administrators.

2. Qualitative and quantitative data collection. Patton (1980) stresses that multiple methods of data collection is fundamental to the holistic design. This includes analyses of quantitative data, questionnaire results, standardized tests, interviewing, and innovative methods such as indicators of performance. "Multiple methods and triangulation of observations contributes to methodological rigor" (Patton, 1980., p. 18). Multiple methods and triangulation refers to the process by which the same object is seen from a

variety of different points of view, based on the assumption that its nature will be better defined and communicated through this process. There are four types: (1) data triangulation--the use of a variety of data sources in an evaluation; (2) investigator triangulation--the use of several different people as the source of information; (3) theory triangulation--the use of multiple perspectives to provide interpretations on the same set of data; and (4) methodological triangulation--the use of a variety of methods to examine a single object, or program (Patton, 1980).

Moran (1987) also suggests a mix of methods. He suggests qualitative evaluation to provide managers with data on constituent perceptions and opinions on the efficacy of procedures and outcomes within the education system. This suggestion counteracts the potential that goals for educational programs are insufficiently designed, or that unanticipated outcomes might arise through too close an adherence to stated objectives. Moran (1987) also suggests the collection of quantitative data for the evaluation. Data established through these procedures are used to check on the degree to which the program is affecting clientele perspective. The result is the development of a process that draws on the strengths of both qualitative and quantitative methodologies.

3. The holistic design versus evaluation research. Holistic evaluation must be distinguished from evaluation research. In true research, a particular ontological belief must guide the study; indeed, much debate rages over which of two methodologies, the qualitative and quantitative, is the correct source of knowledge (Smith and Hesushius, 1987). In holistic evaluation, ontological integrity need not be paramount; what is paramount is gathering the information of necessity to the evaluation-users. Another distinction relates to the issues of internal and external validity (Stufflebeam & Webster, 1988). Research stresses the importance of external validity in a study, so new knowledge--knowledge that can be transferred to other settings--can be developed within the larger educational community. Holistic evaluation stresses internal validity; that is, they are "short term, concerned with extending and generalizing basic understandings about educational processes" (p. 572) within a unique context. They are constructed to meet the needs of the immediate audience, not the larger academic community.

4. Holistic design and stakeholder participation. Cronbach (1980), Moran (1987), Stake (1976), and Stufflebeam and Webster (1988) argue convincingly that involving those affected by the evaluation in the process itself can assist in creating a more effective evaluation. The last state that recent "developments labeled 'responsive evaluation'...are leading in the direction of evaluations that promote ongoing communication and collaboration between evaluators and clients" (p. 573). Such an approach was also suggested as preferential in conducting naturalistic evaluations, and is certainly consistent with the concept of participatory decision making in administration.

Ayers (1987) describes two such approaches as stakeholder based evaluation, and stakeholder collaborative evaluation. The "stakeholder collaborative" (SCE) approach differs from the "stakeholder based approach" (SBE), in that in the SBE the stakeholders provide input to an evaluation which is essentially the responsibility of the evaluator. In SCE, the evaluator and a single group of stakeholders share joint responsibility for the study and resulting reports, and it is the team of participants that is primarily accountable for the report's contents. A comparison of the two approaches is found in Table 2.5.

Stakeholder-Collaborative	Stakeholder-Based
Evaluator and concerned group representatives jointly plan, administer and report results of an evaluation, for which the stakeholder group is viewed as being primarily responsible.	Concerned groups provide evaluation planning and report review input to an evaluation study, for which an evaluation professional is responsible
Characteristics	
stakeholder control	vs. stakeholder influence
evaluator influence	vs. evaluator control
evaluation study	vs. evaluation research
primarily formative	vs. attempt at summative
small-to-midscale	vs. larger scale
internal	vs. external
Predicted Results	
higher acceptance	vs. lower acceptance
greater stakeholder skills result	vs. fewer stakeholder skills result
greater utilization	vs. lower level of use
political pressures on stakeholders	vs. potentially fewer political pressures

Table 2.5: Comparison of "Typical" SBE and SCE approaches (Taken from Ayers, 1987, p. 266).

Table 2.6 delineates the positive and negative features of the SCE approach:

POSITIVE AND NEGATIVE FEATURES OF THE SCE APPROACH		POSITIVE AND NEGATIVE FEATURES OF THE SCE APPROACH	
Positive Features	Negative Features	Positive Features	Negative Features
Different participants representing broad spectrum of view points gives better quality information	Members may not be truly representative	Formative focus of study on improvement of system	Stakeholder group's questions at times not well considered
Provides learning experience for stakeholders, skills that have since been reused	Members may not commit to task and dropout/changeover occurs	More bodies could perform necessary work more easily	Timing of study continuity a problem during summer
Allowed stakeholders to provide input and gave them decisionmaking power	Lack of time, all members had time-consuming jobs apart from the study	Simple, easily understood format of final report	Committees can be difficult to work with
Increased idea-generating brainstorming capability, positive group dynamics	Member burnout, length of study, core-group frustration with non-working members	Increased evaluator credibility as technical resource and "reality tester"	Decreased evaluator autonomy and objectivity, probably a less rigorous study
Stakeholder members "bought into" and expended great effort on behalf of study	Requires a great deal of work and long-term commitment	Overt, public support by district administration for process and resulting report	Possibility of political interference and abuse
Positive response, cooperation from those contacted by group members	Large, complex program to evaluate	Union-management representative cooperation with stakeholder group	Logistical problems: meeting times, locations, secretarial assistance, etc.
	Initially, group roles were not clearly defined or understood by members	Possible increased utilization	Decreased generalizable long-term knowledge

Table 2.6: Positive and negative features of the SCE approach (taken from Ayers, 1987, p. 227).

5. Holistic design and key indicators of performance. Wholey (1983) suggested the use of indicators of performance for improved management. This has become a recent occurrence in the administration of public service institutions such as education (Rossi & Gilmartin, 1980; Johnstone, 1987).

Key indicators (David, 1988; Hanushek, 1986; Murnane, 1987; Murnane & Pauly 1988; Porter, 1988; Richards, 1988; Smith, M. S., 1988; Selden, 1988) are subjective measures of system performance converted into quantifiable units. For example, if one makes the subjective determination that graduation rates are a measure of the system's achievement, then this success measure, although subjective, can be quantified and used as a performance measure.

A fair degree of controversy surrounds the use of key indicators to monitor the performance of education systems. The first source relates to the choice of indicators themselves. How can education performance, particularly with respect to development of a student's character, or self-esteem, be converted into measurable units? As David (1988) states, "Their potential to play a constructive role depends on what they measure and how they are used" (p. 499). Similarly, questions as to indicator design are important. Two questions are important here: (1) what are the characteristics of a good indicator

(Richards, 1988); and (2) who should design them (Porter, 1988): will it be the politicians, who have a tendency to respond to public pressure and only address those issues of public concern, or educators themselves who know intimately the nature of the system?

A third argument centers around their use: will they be used as a mechanism of control, or as a source of information upon which logical rational planning for the future can be based (Porter, 1988)? And, finally, a fourth issue: can we guarantee that we, as educators, who may now be held accountable for the level of these indicators, actually possess the ability to influence scores? Murnane and Pauly (1988) stress, for example, that these measurable indicators of educational performance can have a great influence on policy, and not necessarily a positive one. They state that experience with indicators in economics suggests three lessons for educators: (1) it is important to develop multiple indicators; (2) good indicators make us aware of new and puzzling questions; and (3) it is important to become knowledgeable users of indicators--learn how to design them, measure them, track them, and utilize them for system improvement.

Despite these arguments, a concentrated effort is being made to develop key indicators for education systems. The IPS was originally designed to include key indicators as measures of system performance. A methodological goal of the study is to design a process by which credible key indicators can be developed to characterize the success of the district's mission statement, and that are acceptable to both the educators and the public.

Factors Limiting the Use of Holistic Evaluation Methods

Holistic evaluation is a process that is designed to assist senior administrators to formatively evaluate the needs of their education organization. Yet few school districts have adopted this methodology:

...This lack of interest in holistic evaluation research has less to do with methodological concerns than with the demands these methodologies place on the resources of the organization. These projects involve a comprehensive evaluation of the organization and the social and political environment in which the organization operates. This generally involves significant expenditures of time, effort, and money that few organizations are willing or able to make. Additionally, few public officials have the requisite research background to appreciate and meaningfully participate in this type of endeavour. Furthermore,

many researchers do not possess the requisite management background to define appropriate research questions for this type of activity (Moran, p. 613).

For holistic evaluation to be employed on a systematic basis, a massive introductory training program, supported by significant resource allocations to assist in its implementation, must be forthcoming. Stufflebeam and Webster (1988) make a number of recommendations in this regard. A first suggestion is that administrator pre-training programs should institute courses in program evaluation--or provide them with field experiences in evaluation. A second suggestion is to establish in-service training for existing administrators. Thirdly, they suggest most districts do not have a comprehensive evaluation policy, and therefore have not developed their evaluation capacities; they should do so. And finally, they suggest that evaluations that are being done should meet accepted standards of evaluation practice.

Standards of Practice for Program Evaluations

The Joint Committee on Standards for Educational Evaluation outlines the standards established for program evaluations in the United States (Stufflebeam, 1989). They are apropos to holistic evaluation because it is an evaluation process of a compilation of programs; also, they parallel to a great extent some of the qualities expected of effective administration. The four standards are:

1. **Utility:** Sound evaluations deliver information that is useful to the audience it is intended to serve (suggests a stakeholder approach for holistic evaluations of districts).
2. **Feasibility:** Feasible evaluation procedures are ones that can be carried out in the real world (the reason for the action research component of the study, and the reliance on stakeholder opinions as to its effectiveness).
3. **Propriety:** Proper evaluations are ethical (consistent with the expectations of effective administration).
4. **Accuracy:** Sound evaluations produce accurate information.

Each of the standards has criteria that more clearly articulate expectations for program evaluations. These are described in Table 2.7. Many of the specific

expectations have been discussed at length in this chapter, suffice it to say that efforts were made during the action research component of the study to adhere as much as possible to these standards..

Table 2.7: Standards of good practice for Program Evaluation (Stufflebeam, 1989).

1. Utility:

- a. identify audience and consult with members to determine questions to be asked;
- b. Staff the evaluation with persons qualified and acceptable to the audience.
- c. collect an appropriate range of information; be selective so as to address the most important questions.
- d. describe the perspectives procedures, and rationale used to interpret findings, so that the valuational interpretations are clear and well documented.
- e. Clearly report the object of the evaluation, its context, the findings and the evaluation procedures.
- f. Make sure that evaluation reports are disseminated to all the relevant audiences.
- g. Schedule data collecting and reporting activities so that reports are delivered when needed.
- h. Help with follow-up work and promote utilization of the findings.

2. Feasibility:

- a. Choose and implement practical procedures that minimize disruption to the educational processes being evaluated and that are feasible and reasonable given the constraints of time, budget, staff, and data availability.
 - b. Obtain the cooperation and support of the interested parties: sometimes political viability of an evaluation can be aided by organizing and involving a representative advisory group.
 - c. Be efficient in the use of money and people's time; do all you can to make the evaluation cost effective.
-

Table 2.7: (cont.) Standards of good practice for Program Evaluation

3. Propriety:

- a. Negotiate a clear written contract or memorandum of agreement to govern the evaluation.
- b. Identify potential conflicts of interest and deal with them openly and honestly.
- c. Exercise full and frank disclosure of findings and limitations of the evaluation.
- d. Release the findings to all parties who have a legal and ethical right to know them.
- e. Protect the individual rights of parties to the evaluation.
- f. Exercise appropriate human interactions by being respectful and considerate of the parties to the evaluation.
- g. Give a balanced report of the program's strengths and weaknesses.
- h. Exercise fiscal responsibility in using and accounting for resources.

Accuracy:

- a. Carefully describe the object of the evaluation.
 - b. Examine and report the context of the program in enough detail.
 - c. Describe both the planned and the actual purposes and procedures of the evaluation.
 - d. Describe information sources in enough detail, so that the adequacy of the information can be assessed.
 - e. Choose or develop evaluation instruments and procedures so that the interpretations arrived at are valid for the given use.
 - f. Establish and document measurement reliability, so that findings are consistent and dependable for the intended use.
 - g. Exercise systematic data control.
 - h. Appropriately and systematically analyze quantitative information to ensure supportable interpretations.
-

Summary

This chapter has concentrated on outlining the processes of administration and evaluation; their purposes, various models and strategies, the factors both enhancing and limiting their effectiveness; and the appropriateness of holistic and naturalistic evaluation as an administrative tool for school district evaluation. From these deliberations direction for the overall design of the study, suggestions for improvement of the IPS model itself, and criteria for determining effectiveness of the System can be drawn.

The suggestions were employed as much as possible during the action research part of the study in order to improve the IPS's methodological integrity. In addition, there appears to be three main groups of factors that affect the effectiveness of evaluation as a leadership substitute: (1) methodological procedures; (2) administrative acceptance and actions during its implementation; and (3) resource allocation and expertise of administrators. Insights into how these factors interact so as to limit the effectiveness of the IPS will form a major part of the findings in the study. Finally, what implications these findings have for the validity of the leader-substitute construct of leadership itself will conclude the investigation. Details of the study's design are presented in the next chapter.

CHAPTER III

Design and Methodology

Action Research and Naturalistic Inquiry

It was suggested in Chapter II that the IPS model used in 12 school districts in the 1986-87 school year needed significant refinement in order to meet its objectives. In order to achieve this refinement, it was decided to run a second series of pilots and incorporate these as action research. A naturalistic evaluation of the action research was then done to achieve the remainder of the study's purposes. To ensure the action research and naturalistic evaluation was properly grounded in previous research the literature search chronicled in Chapter II was conducted and then taken into account.

Kemmis & McTaggart (1988) describe action research as:

...a form of collective self-reflective enquiry undertaken by participants in social situations in order to improve the rationality and justice of their own social or educational practices, as well as their understanding of these practices and situations in which these practices are carried out. Groups of participants can be teachers, students, principals, parents and other community members,--any group with a shared concern. The approach is only action research when it is collaborative, though it is important to realise that the action research of the group is achieved through the critically examined action of individual group members. In education, action research has been employed in...school improvement programs and systems planning and policy development...(p. 5)

The action research was conducted according to the methodology and form of a naturalistic evaluation (Lincoln & Guba, 1985). Such an approach is consistent with action research and the emergent design of the IPS model, and is suggested when in-depth understanding of a concept is the intent of a study. In the particular instance of the IPS evaluation model the nature of the problem studied (i.e., determining limitations of the IPS while at the same time attempting to overcome them) was not fully determined at the outset. As such, the change in the problem was anticipated as the inquiry proceeded. Therefore, the design could not be fully specified in advance but was expected to emerge over time, even though sufficient grounding (i.e., a pilot project

that was a partial success, and theoretical support for the IPS) was in place for the research described below. Figure 3.1 provides a summary of the flow of naturalistic inquiry as planned for this study.

Data collection and data analysis went on contingently and virtually simultaneously. Data collection methods were altered and adjusted according to individual circumstances in each case study. Data interpretation was a function of the natural setting of each school district. Presentation of this information is in the form of a "thick" case study. These principles are commensurate with the requirements for a naturalistic study as presented in Chapter II. The study concentrates on use of the IPS model in three school districts. Some potential procedural concerns attendant to the naturalistic design will be discussed before outlining the procedures themselves.

Action research employing the IPS evaluation model was conducted in seven school districts during the 1987-88 school year. A naturalistic meta-evaluation of its results took place during and after completion of the district evaluations. Three of the seven districts were chosen as the focus for the meta-evaluation. The action research then became a source of data to seek out answers to the central problem of this study, and its three subsidiary purposes.

The Problem

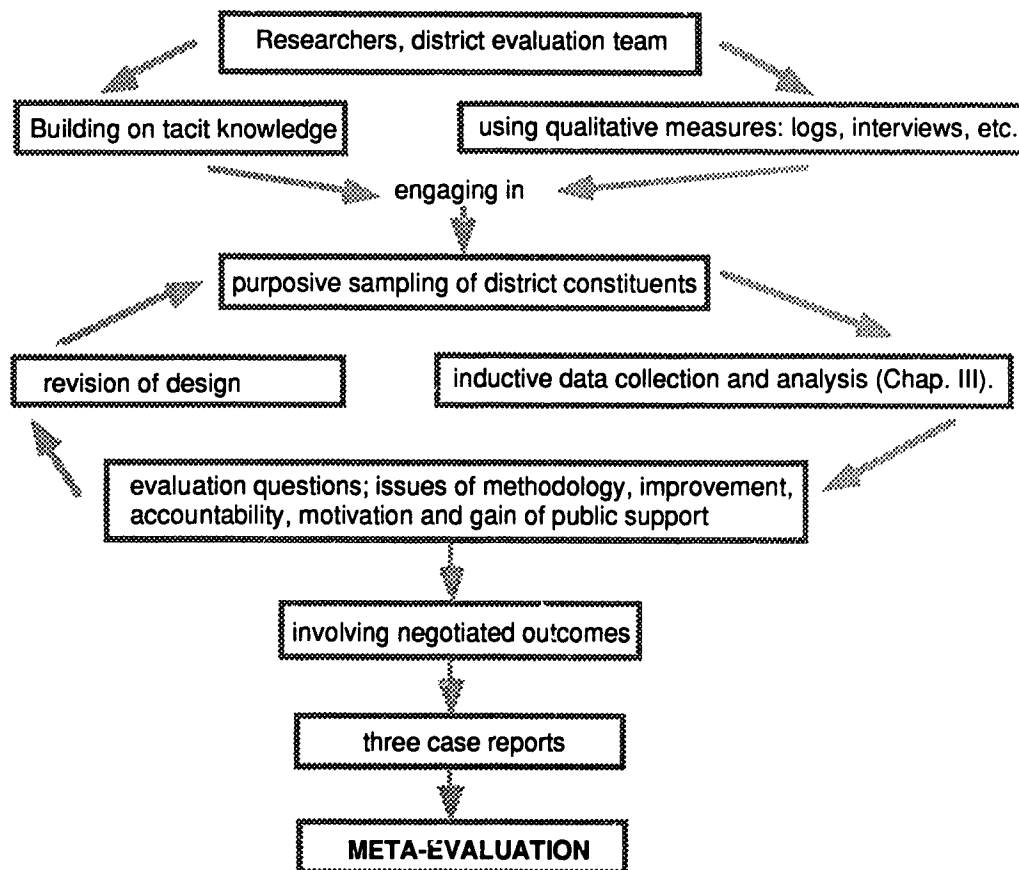
The central problem was to ascertain whether the specific IPS evaluation model could be refined and reformatted to act as an effective leader-substitute for educational administrators in British Columbia. Examination of documentation of the first pilot and of literature on administration and the leader-substitute construct (Chapter II) suggested (1) that administrative style, ability and skill is enhanced or neutralized by its interaction within a context; and (2) that organizational evaluation is a strategy that can, theoretically at least, lead to achievement of administrative purposes such as vision, management of meaning, and empowerment. However, examination of the literature on evaluation indicated that evaluation has only had limited success as an administrative tool. If one relates the two literatures, it appeared that a possible reason might be due to the fact that evaluation procedures and

Table 3.1: The Naturalistic Design Applied to the Meta-evaluation Study (Note: adapted from the book Naturalistic Inquiry by Y. Lincoln & E. Guba (1985), p. 188.)

PROBLEM:

- (A) To refine and critique IPS
- (B) To understand the limitations of formative evaluation in school districts
- (C) To gain insights into formative evaluation, administration, and the leader-substitute construct

Natural setting: Districts of Inlet, South Coast and Interior.



processes need to complement the nature of administration in a holistic organization. This in turn suggested that the IPS model should be a holistic evaluation process in which tasks, structures and subordinate characteristics (i.e., leadership skills, value commitments, etc.) were used to enhance the administrative capability. The central problem of the study became, therefore, to develop a system that enhanced the leadership skills of vision, empowerment, and management of meaning so as to achieve the IPS model's original purposes: improved accountability, professional teamwork and commitment, improved decision making, and public support for education.

The central problem was ultimately resolved into three sub-problems:

1. Refinement and Critique of the IPS.

A first aim was to discover better methodologies and implementation strategies for the IPS model in each school district. This was the focus of the action research, in which the researcher attempted to adjust practice while assisting with the three individual evaluation processes. Also, a critique of the procedures and processes used in each district evaluation, using the recorded empirical observations and experiences of stakeholders, was conducted to refine the structure and process of the IPS. A meta-evaluation of the experiences of the three cases was the final source of information for the System's refinement.

A meta-evaluation, in a summative sense, of the overall effectiveness of the IPS model as a administrative tool for educational leaders was also conducted. Effectiveness was measured in the domains of accountability, teamwork and commitment, decision making, and public support for education.

2. To Understand the Limitations of Formative Evaluation in School Districts

A second aim was to gain further insight into the contention that while formative evaluation seems to hold great promise as a tool for system improvement, it rarely is as effective as might be supposed. The same meta-evaluation was to be used as an empirical data source from which to gain insights into the factors that limited the use of formative evaluation as a tool for school district improvement.

3. Formative Evaluation, Administration, and the Leader-substitute Construct.

A final aim was to examine more closely the interaction between formative evaluation and administration, in order to understand the true potential of formative evaluation in administration. In addition, the data and findings of the research was also studied so as to gain a better understanding of the leader-substitute construct.

Strategy and Procedures

The study was conducted in a series of four stages, or steps. The first stage was to conduct action research in each of the three districts and to chronicle the efforts made to expand, refine, and improve the evaluative and administrative procedures of the IPS. The second stage was to observe, document, and evaluate the overall success of the district evaluations, with respect to the goals established for the system by its originators: better accountability, improved professional teamwork and commitment, improved decision making, and maintenance and gain of support from both the public in the district. The third stage was to examine the factors that limited the achievement of these goals, and to relate them to what is known about formative evaluation and administration. The final stage was to examine the implications of these findings with regard to the leader-substitute construct of leadership.

1. Refining the IPS

Action research was the first strategy employed to achieve the purpose of improving the IPS model. Kemmis & McTaggart (1988) explain the reasons why this approach was deemed suitable in the circumstances surrounding use of the IPS :

Action research can be seen as an approach for groups of educational practitioners, students' parents' and others to live with the complexity of real experience while at the same time, striving for concrete improvement. It is a way of managing complex situations critically and practically (p. 7).

Each of the districts wishing to employ the IPS in the 1987-88 school year also wanted the opportunity to adapt it to their unique needs and circumstances. This was done by an evaluation team composed of educational leaders in each

district. Therefore, the action research strategy provided a process whereby improvement of the model could be collaboratively pursued.

To adjust the overall conceptual model of the IPS a meta-evaluation of the experiences of its use in all three districts was also conducted. Four questions derived from the previous pilot (outlined in Chapter II) guided this section of the study. They are:

1. How acceptable is the Four-goal Six-attribute framework in characterizing the overall purposes and methods of the education system?
2. What indicators, other than those currently in use, are appropriate measures for each goal and attribute?
3. What changes should be made to the questionnaires and reporting-out formats to make the IPS more effective? and
4. What processes and procedures were valuable in focusing evaluation team members on the task of evaluation?

To answer these questions, stakeholder collaborative evaluation teams were established in each pilot district. These teams were charged with the responsibility to seek out and define appropriate key indicators for each goal and attribute, to analyse and refine the four-goal and six-attribute framework, and to create and publish an evaluation report. A final task for this group was to comment, in a critical fashion, on the efficacy of the IPS model's procedures and processes. Stakeholder team members were interviewed and their comments used to help refine the process and procedures of the IPS model.

The interview protocol asked stakeholders about the four questions shown above (see Appendix B1). Of particular interest were queries relating to: (1) the suitability of the process for determining key indicators; (2) the success of different survey/questionnaire techniques, (3) the efficacy of data collection, interpretation, and analysis techniques; (4) the appropriateness of the values framework; (5) the efficacy and integrity of the evaluation report, and (6) the overall efficiency and effectiveness of the IPS model.

Minutes of team meetings and records of the deliberations and decisions were kept either by an official recorder or the researcher. A comprehensive list of indicators was compiled and subjected to reference checks by other members of the educational community. Qualitative decisions as to their

credibility were made accordingly. A further data source was the documentation undertaken by individuals involved in the evaluation in the three districts. Stakeholders were asked to submit suggestions for improvement of the process, and senior administrators provided letters and information files that they compiled while undertaking the process. This data was analysed for suggestions for improvement to the IPS.

2. Determining the Effectiveness of the IPS

To estimate the degree of the IPS 's effectiveness, measures of the four goals established for it by its progenitors was undertaken. These are: (a) improved accountability; (b) improved sense of teamwork amongst educational leaders; (c) improved decision making; and (d) maintenance and gain of public support for education.

(a) accountability. A first determination of accountability was to measure the effort made on behalf of the district to gather evaluation information, and to distribute results to the various constituents. Procedures to accomplish the first of these two goals were built into the evaluation, and a district that completes the evaluation is therefore deemed to be partially successful in this regard.

A second indicator of accountability was to measure the effect of efforts made by the district to distribute evaluation information, and to engage in follow up discussion to the evaluation activity. Greater accountability leads to greater understanding about education; people feel knowledgeable about the nature of the education system they are working in. Therefore, stakeholders were asked about the educative quality of the IPS. Efforts made by the district in the first three months following the publication of the evaluation report to distribute it and to create discussion of its contents were also analysed. In addition questions as to its effectiveness were asked of evaluation team members and parents, other district people, and teachers.

(b) Teamwork and commitment. Evaluation can accomplish both complacency reduction and conflict resolution, and it can be carried out in such a way as to create a greater sense of teamwork and commitment from various groups within the educational community. The literature review suggested that evaluation processes that are perceived to be ethical, that are empowering, and that are participatory contribute to the creation of teamwork

and commitment. How judgments were made as to the success of the IPS in achieving these characteristics is explained below.

(1) Questions relating to these characteristics were asked in the interview protocol. They are adapted from a questionnaire developed by Ayers, (1987). Stakeholders were asked whether it was a rational, honest, open, and just process with integrity; they were also asked whether "authority commensurate with responsibility" was created in the evaluation. Other questions asked were: "were politics conducted responsibly", "was the evaluation dignifying of constituent opinions", and "was excessive executive authority used?" These questions were adapted from an instrument designed by McPherson (1988).

(2) In the same interviews role descriptions of each stakeholder were examined for evidence as to whether the process engendered greater feelings of personal control, willingness to work, and a sense of accomplishment. This was done through the process of semantic differentiation, a process adapted from an interview protocol designed by Ayers (1987). Responses to other questions were content-analyzed to seek out further evidence.

(3) Stakeholders were asked to comment on the extent to which they themselves, and their constituencies, were willing to participate in future district improvement activities (i.e., attending meetings, assuming responsible roles, initiating new projects).

(4) A number of community and parent meetings were attended and participants interviewed to find out whether they perceived themselves to be more informed or more committed as a result of participating in the follow-up to evaluation. Observations were documented and analyzed.

(5) If conflict was manifest in "group" disagreement-- i.e., between the teachers and school trustees--assessments of cooperation before and after the evaluation were made. Estimates as to the strength and degree of conflict, based on the conduct of group representatives were also determined.

(6) An assessment of the ability of the evaluation process to accomplish administrative goals such as vision-setting and information management was done during the interviews of evaluation team members (see Appendix B.1 for this instrument). The purpose of the interview was to see if those aspects of leader behavior leading to better conflict management have improved as a result of the use of the IPS.

(7) A review of the newspaper articles on education--for the three months after the use of the IPS--was done. Stories were analyzed for insight into the contribution of the IPS towards either greater cooperation or greater conflict within the district.

(c) Decision making. Judgments as to whether decision making had improved or not was based on the following criteria:

(1) Stakeholder perceptions as to whether a clearer direction for the district, or a clearer focus as to the parameters surrounding decision making resulted from the evaluation.

(2) Whether or not a long range plan was developed, as a consequence of the IPS's application. Did the administration utilize the report in a consultative nature to develop decisions, or did it not?

To make these determinations, three procedures were followed. First, the researcher or his assistants attended as many planning meetings as possible, and recorded observations as to the progress towards planning. Second, examinations of district reports to both the school board and public were conducted. They were examined for statements indicative of planning. Third, stakeholders were asked for knowledge of where evaluation information resulting from the use of the IPS model was used in decision making. A district that developed a long range plan as a result of the evaluation was deemed to have demonstrated improved decision making; a district that used the evaluation information either formally or informally to inform decisions only partially improved decision making.

To assist in data collection, the researcher or assistant kept a log of all meetings and deliberations connected with the evaluation. The logs will be used to present the context within which these decisions were made, and to provide the "thick" nature of the case study explanation so that the dynamics of the decision making process can be understood.

(d) Public support. Success in achievement of this goal will be measured indirectly. First, the district strategy for applications of the evaluation information was examined in terms of its comprehensiveness and potential effect on public support. Second, the Stakeholder Team members were asked to evaluate the actual and potential effectiveness of the strategy on public support.

A third method was to attend representative meetings of parent groups to observe the effect of the process, and to gauge the degree of support for education as a result of the evaluation process. Observations and questionnaires used at parent meetings will also be analyzed for information relating to this goal.

3. Formative Evaluation's Limitations and the Leader-substitute Construct

The ultimate result of all the efforts made in each case study to determine evaluation methodology and effectiveness was to provide the researcher with experiences and empirical evidence that could lead to better understanding of how formative evaluation, at the school district level, functioned as an administrative tool. In addition, insights into the leader-substitute construct of leadership were sought.

Observations of stakeholder conduct during the action research were a fundamental data source for these two topics. Further insights were derived from analysis of both formal and informal interviews with stakeholders. Two separate debriefing sessions were held with district coordinators for the IPS project in order to get a complete picture as to what was done as a result of the evaluation and to determine the "limitations" as to its success. During the last case study interviewees were asked to prioritize factors that influenced them the most during the evaluation, factors that reflected the concepts inherent in the leader-substitute construct. Judgments as to the realistic potential for the use of formative evaluation by administrators for improvement of the educational organization were made on the basis of analysis of all this information.

Sample

The decision to delimit the study to only three of seven districts, as opposed to a larger sample, was made to enhance the contact between the Ministry participants and the districts being evaluated (Dickson, 1988b). A smaller sample was recommended because this allowed the researcher to spend more time in each district and therefore to experience firsthand the problems with the evaluation model. Fine tuning is only possible when an analysis of district needs and purposes is done in a detailed and in-depth fashion. The assumption here (as in all naturalistic studies) is that the complexity and individuality of each context is so unique that only through true in-depth analysis can an

understanding of the dynamics affecting administrative action be understood. The districts chosen to be in the study were taken from a potential sample of seven that volunteered to use the IPS model to evaluate their educational programs in the 1987-88 school year. These seven districts were representative of six of the eight different categories into which the 75 school districts are placed by the Ministry of Education for comparative purposes (Cameron, 1987). The three that were selected were chosen for the following reasons:

1. Each district was similar to a significant number of other districts, so as to maximize the potential for transferability;
2. Each district was reasonably accessible so as to ensure prolonged engagement and the opportunity for regular contact; and
3. Each district was prepared to commit the resources and manpower to the action component part of the research, and was prepared to allow its efforts to be meta-evaluated and subsequently published.

Participating districts in the sample, from smallest to largest, were School District #1 (Inlet), School District #2 (South Coast), and School District #3 (Interior). The demographic characteristics for each district are outlined in Table 3.2. The districts range in size from approximately 15000 students (Interior), to approximately 6500 (South Coast). Interior can be described as a fairly large mixed urban and rural district with a mix of large and medium-sized schools. Inlet is a medium-sized district with an urban center, but removed from the province's main population centres; it also has secondary population areas that are distinct from the urban center. It has a mix of large and small schools. South Coast is also a medium-sized district, but is dominated by its suburban area. It has a large range of school size as well, with two community secondary schools. Further details will be provided in each case study.

Data Collection Techniques

Data utilized in the study were gathered in a variety of ways, principally by means of interview. The interview's design and administration was guided by requirements outlined by Spradley (1979). Each stakeholder was deemed to be a key informant (LeCompte & Goetz, 1980). Interview questions were derived

Table 3.2: Demographic characteristics of Inlet, South Coast, and Interior School Districts (District Information Sheets, 1988: PGE & R, BCME).

	Inlet School District	South Coast School District	Interior School District
F.T.E. enrolment	6,043	7,454	14,965
Assessed Value/ Pupil as represented by property tax base	\$82, 815	\$128,628	\$159,200
Average Income of Household	\$32,352	\$25,726	\$29,493
1988 Tax paid by half of home owners after application of B.C. Home owner grant*	\$-30.00	\$219.00	\$59.00
Native Peoples as % of total pop.	7.8	1.1	0.9
Non-English Home Language as % of total pop.	6.2	1.8	5.6
Children of Lone Parent Families as % of all children	16.9	15.0	18.2
High school graduation or better as % of total pop.	45.9	43.9	44.8

* Government subsidy provided to home-owners to reduce actual property tax paid for schools.

from the expressed purposes of the study and were chosen to reflect the pertinent concepts suggested by the literature review. In most instances the questions or procedures were adapted from interview protocols used previously in evaluation studies by Ayers (1987) and McPherson (1988). Closed and open interview questions were asked; and both quantitative and qualitative responses requested. Interview times varied between one hour and an hour and a half.

Interviewees were contacted well in advance of the interview, and interviews were held in school board offices in the district. All but two interviews took place in a quiet, private setting. In the case of the two exceptions, the tone of the interview was slightly altered due to the interruptions. In most instances the interviewees were described as relaxed or positive; on one occasion (with the parent in Interior), the interviewee was described as nervous and shy; on a second (the Director in South Coast) the interviewee was described as distracted, and somewhat inattentive.

To establish a climate of trust, and to get to know interviewees, an informal interview with most stakeholders was done early on during the action research component in both the Inlet and South Coast districts. These notes were also deemed to be part of the data for the naturalistic evaluation. A similar interview was not conducted in the case of the Interior evaluation because it was not deemed necessary.

Interviews were held (in the case of the Inlet and South Coast districts) three months after publication of the evaluation report. Due to feedback that this was too long a period to wait, interviews in the Interior District were conducted two months after the completion of the report. Whereas this was helpful in getting feedback on the evaluation process itself, it required a follow-up contact to gather accurate answers to questions on applications of the evaluation information.

Interviews were taped and rough notes were taken. The researcher transcribed each interview and placed each individual question/response on a card. These cards were then content analysed and color-coded so as to be categorized according to which purpose of the study they related to. They were then grouped by color, and re-analyzed.

Other data were compiled from the logs compiled for each district evaluation by the researcher or his assistant. These consisted of meeting

minutes, observations, critiques, and reference checks with one another. Data within these logs were itemized, placed on cards, color-coded, and added to the data from the interviews. So was any pertinent information derived from analysis of rough notes from the meetings, district documents, or informal interviews. Patterns within the data characterizing each district's evaluation were then sought out.

From these patterns the conceptual framework for the case studies themselves was developed, and answers for questions relating to the meta-evaluation were created. When seeking patterns across all three districts for the meta-evaluation component of the study, a similar process was used. In addition, the quantitative ratings on questions were compiled for each district and for comparisons across districts (these are found in Appendices F1-F12).

As a result of all these activities, the researcher analyzed the experiences and data so recorded to "meta-evaluate" the effectiveness of the IPS in light of its intended objectives. This analysis attempted to adhere to the processes described below, in which methods as to how to generalize from multiple case studies is discussed.

The Case Study Approach

Lincoln and Guba (1985), and Alkin, Daillak and White (1979) suggest that the case study is an ideal method for the presentation of data in a naturalistic evaluation. Lincoln and Guba (1985) define it as "'a slice of life' or a 'depth examination of an instance'" (p. 360), or as they quote Denny (1978), an "intensive or complete examination of a facet, an issue, or perhaps the events of a geographic setting over time" (p. 360). Case studies will vary (1) from setting to setting; (2) in purpose; (3) in level of specificity, and (4) in level of analysis--i.e., from factual to interpretive to evaluative.

Case studies are suitable to naturalistic evaluations for a variety of reasons. Case studies contribute to the development of understanding, from both the point of view of writer and reader. From the writer's viewpoint, case studies can be produced in harmony with the evaluator himself, flowing from his tacit knowledge and understanding of the situation. Secondly, they allow for in depth or "thick" analysis; not just a chronology of events, but also the meanings and interpretations of events--thus the essence, or chemistry of the event--its true meaning, can be characterized. Thirdly, the case study is an

outgrowth of the inductive and continuous nature of naturalistic inquiry--whereby grounded theory is derived out of constantly interpreted experience. Case studies also allow for reader interpretation, building on the reader's tacit knowledge for understanding. It also provides the reader with context information, so transferability can be judged.

Case studies should contain the following elements (Lincoln & Guba, 1985):

1. an explanation of the evaluand (i.e., the object being evaluated); in this case, the school district and its evaluation program;
2. a thorough description of the context or setting within which the evaluation took place and with which the study was concerned;
3. a thorough description of the transactions or processes observed in context that are relevant to the evaluation; that is, those elements identified as important are studied indepth;
4. an explanation of the outcomes of the study, that is, the results of the evaluation.
5. a discussion of the important factors identified at the site as salient to the outcomes of the study: i.e., why the evaluation was effective or not effective (that is, methodologically and circumstantially speaking).

Substantive considerations (i.e., those mentioned above) and methodological considerations--that is, the steps taken in terms of method, investigating, and establishing trustworthiness--must both be part of the case-study report.

1. Generalizing from multiple case studies. Greene and David (1984) suggest a design for generalizing to a target population of cases from the results of a purposefully selected sample of cases. The four main features of such a design are:

- (a) a conceptual framework that provides the superordinate structure;
- (b) a sampling plan that ensures representativeness of the target population in the sample of cases;
- (c) procedures for the conduct of individual case studies that insure sufficient comparability across cases;
- and (d) a cross-site analysis strategy that tests the limiting conditions of the findings (p. 75).

The conceptual framework should limit and focus the scope of the study, by identifying the main elements and events of interest, and the main features of

the context in which these factors are operating. In this study a conceptual framework was established that travelled to each venue: i.e., the use of a "values" framework; the use of key indicators; the use of a stakeholder collaborative team; general methodology; and reporting out procedures. In addition, a design has been developed that focuses questions on two main areas of interest: methodological improvement, and overall effectiveness (i.e., in terms of utilization and influence). Within this general framework naturalistic inquiry was taking place.

The sampling for this project was described earlier in this chapter; suffice it to say that no efforts to generalize beyond the representativeness of the purposeful sample can be attempted; and no presumption of applicability except as perceived by the interpreter should be made, in keeping with naturalistic precepts re transferability and generalization.

To implement a multiple case study design to allow for generalizability of findings one must also ensure that data collectors think and act more or less alike. In this circumstance the three interviewers were trained to accomplish this objective. Interview guides were used, interview practice was initiated, and interviewers visited the sites and participated in all other related activities associated with data collection. Debriefing and cross-referencing was also utilized.

Validity and Reliability

Trustworthiness

Trustworthiness refers to the qualities of credibility (internal validity), transferability (external validity), dependability (reliability) and confirmability (objectivity) (Lincoln and Guba, 1985). These must be established for each case study to ensure the academic integrity of the study. Some of these requirements were described in Chapter IV. A first requirement is to establish an audit trail. Appendix B.2 outlines the audit trail for this study.

A number of writers have indicated that certain practices, undertaken during a naturalistic evaluation (as the meta-evaluation is in this case), maximize its potential to support the quality of trustworthiness. Greene (1987a) suggests that "a successful practical naturalistic evaluation is contingent upon meaningful participation of significant user groups in the evaluation process and on the open expression and legitimation of the multiple

value claims held by members of such groups" (p. 329). The procedures outlined in this study are intended to ensure multi-participation of clients.

Pearsol (1987) suggests that a key to successful evaluation is "not the meeting of trustworthiness criteria that satisfy the dependability of a set of evaluation conclusions" (p. 340). Rather, it is the evaluator's conscientious attempts to synthesize data and implicit or explicit value claims that will lead to justifiable evaluative conclusions. This approach is built into the methodology of this study.

Joanne Farley (1987) claims that conclusions in naturalistic evaluations are justifiable when they meet three types of acceptability criteria. First, empirical claims must be empirically accurate; that is events did transpire as they are described in the study. Second, the description of events must reflect an accurate reconstruction not only of events, but the meaningfulness of the events. Finally, she states that "the interpretive explanation and related conclusions must be normatively adequate" (Farley, 1987, p. 350). In other words, the explanations given of various decision making processes conducted by participants in the evaluation must encapsulate the essence of the value choices inherent in those decisions.

The suggestions made by Pearsol, Greene, and Farley are partially ensured through employment of a number of techniques suggested by Lincoln and Guba (1985) to ensure credibility. The methods themselves and the efforts made to achieve them in this study are listed below.

(a) prolonged engagement: This means being in contact with the project long enough to take into account distortions that might creep into the data, but not so long to go "native". This was ensured by establishing and maintaining contact with district representatives over a 10 month period, during which time at least six visits to each district were made.

(b) persistent observation: This refers to identifying those characteristics and elements in the situation that are most relevant to the problem or issue being pursued and focussing on them in detail. This was achieved through informal interviews as the project progressed.

(c) triangulation: Triangulation means using multiple and different sources, methods, and investigators (different modes of data collection). In this instance a variety of data sources were used, and two research assistants were employed to assist with investigations.

(d) peer debriefing: This means employing the process of exposing oneself to a disinterested peer in a manner paralleling an analytic session and for the purpose of exploring aspects of the inquiry that might otherwise remain only implicit within the inquirer's mind. This was done regularly through regular contacts between the researcher and his assistants.

(e) negative case analysis: Negative case analysis means continuously refining a hypothesis until it accounts for all known cases without exception; the search for patterns as discussed by Greene and David (1984). Attempts in this regard are revealed by the logic of analysis extending through Chapters IV - VIII of the study, and the data interpretation procedures described earlier.

(f) referential adequacy: This concept refers to the recording of materials so that they provide a benchmark against which later data analyses and interpretations could be tested for adequacy. It means surrendering some of the hard-won data to an archive, for later reference. The material so selected must be representative. Samples of such data are included in the appendices of the study.

(g) member checks: This refers to testing data, analytic categories, and interpretations and conclusions with members of those stakeholding groups from which the data were originally collected. This was done during stakeholder meetings--however, for confidentiality reasons stakeholders were not asked to read over the final case study reports, as quotations contained within could be easily identified with the person making them.

To ensure transferability, detailed descriptions of each pilot context and activity are provided in each case study.

Limitations and Delimitations

Limitations

A first potential design concern is the issue of the Hawthorne Effect (Hanson, 1985). The first of two potential problems here is that the researcher injected himself as facilitator into each district evaluation, and therefore may have influenced the results of the evaluation in either of two ways: (1) as a surrogate administrator, influencing the dynamics of the district context; and (2) as a novelty, creating a positive reaction according to the findings of the Hawthorne group in the classic study (Roethlisberger & Dickson, 1939). Involvement of the researcher was necessary to facilitate the action research

component of the design. The fact that the same factor was found in each case should minimize the effect on the results.

The Hawthorne effect also suggests that an innovation can be successfully employed as a single event because of attention being paid to the individuals participating in it; then the impact tends to wear off as the control of the informal group reasserts itself on people's behavior. Leadership initiatives become "subverted" by the informal relationships in an organization. Successful leadership initiatives should be longlasting; therefore, they should either restructure the informal organization or take advantage of it to embellish the performance of all concerned. The timeline in this study did not allow for measurement of the effect long term use of the IPS would have on school districts, and evidence exists to show that the informal group did reassert itself on people's behavior even within the timeframe of the research.

A factor mitigating the potential confusion of the Hawthorne effect is the naturalistic design itself. It could itself be considered a methodology for creating an understanding of the dynamics of such variables. If the Hawthorne Effect is an inevitable component of any innovation, then it is a factor that all administrators must deal with when using the IPS model. Novelty is also an important component of effective teaching; it is a process by which interest and motivation is encouraged. So if the administrator's purpose is to develop a "learning" organization through formative evaluation, new experiences may in fact be required to engender growth. Since one of the purposes of the study was to refine the system to make it more effective, then this factor is implicit in its reconstruction. Concerns about the potential limitations of the Hawthorne effect can only be minimized through raising it to awareness; it is up to the naturalistic evaluator to make judgements by "factoring out" its apparent effect. To assist in this regard "novelty" was listed as a potential reason for the evaluation's success when asking interviewees in the Interior district the factors that contributed most to its achievements.

A second potential limitation--this time to credibility--is the potential for "evaluphobia." This antipathy to evaluation limits the openness and honesty of participants in the process (Scriven, 1983b). Although it might be assumed that the administrators who volunteered for the evaluation may be relatively free of evaluphobia, the same cannot be said for other members of their

constituencies. To deal with this concern each evaluator will prepare participants prior to participating in the evaluation. Even so, it must be realized that they will differ significantly in their success. Evaluphobia cannot be eliminated, it can only be brought within awareness; and then countered as effectively as possible through the proper use of good interview techniques and through careful discounting and consideration in the evaluation itself.

Delimitations

Delimitations are those factors that deliberately or by happenstance place bounds around the study. In this instance there are three: (1) its timeline; (2) limitations of the IPS itself; and (3) the sample (*q.v. supra*).

The timeline for the study did not extend beyond the first implementation of the IPS. It was originally anticipated that a year and a half timeline would provide all the data necessary to understand the needs and limitations of the IPS model. However, it becomes apparent (see Chapter VIII) that a three year time period, extending into a second use of the model, would be preferable.

It was also discovered throughout the research that the IPS model itself was not really designed as a formative evaluation model, but as an event evaluation (see page 236). Its use for formative evaluation must be therefore be viewed with caution.

**CHAPTER IV:
CASE STUDY NUMBER ONE:**

The Inlet School District Evaluation

Inlet school district is located on the southwest coast of British Columbia. It contains 6200 students and covers approximately 5000 square km. Inlet is considered a medium sized school district in B.C.¹

Economic livelihood is supported by traditional staples of the B.C. economy: forestry, fishing, and tourism. Four main communities predominate. All are distinct in their economic and cultural identity. Linked by an east-west highway, they are within two hours of each other and accessible in three to five hours from Vancouver (the largest city in B.C.) and Victoria, the provincial capital.

In the eyes of its educators, students, and public, Inlet school district is unique--in both its accomplishments and problems. This case study attempts to capture the individuality of Inlet through both the eyes and ears of researchers and participants in the district evaluation. As a conceptual base to this and the other case reports, the efforts made to expand, refine and improve the evaluative and administrative procedures of the IPS will be presented under a series of headings that describe the procedures followed. Each procedure is subsequently described and then critiqued. Information, analysis and discussion will be treated under the following headings:

1. The Context and Background to the Evaluation.
2. Description and Critique of the Procedure and Processes
Followed in the Evaluation.
4. Description and Critique of Applications of the Inlet District Evaluation.

Evaluation and meta-evaluation of this and the other district evaluations will be dealt with in Chapter VII.

¹ British Columbia has 75 school districts ranging from 200 square miles to 33,000 square miles, and from 400 to 50,000 students.

The Context and Background to the Evaluation

Inlet's major city, Cape George--on the eastern extremity of the district--presents two poles of vivid contrast. One is first struck by the immense natural beauty of its setting, a setting soiled and made tawdry by the stench of pulp mills. The smell dominates the environment as does the paper company the economics of Cape George. The pulp and paper industry is a stabilizing force in that its presence assures the town's livelihood; yet being so intimately tied to the forestry industry, it is subject to the vicissitudes of a natural resource economy. Until 1982 Cape George had the second highest per capita income in Canada, but since the 1982 recession, times--and incomes--have suffered. With less wealth, however, has come less transience and stronger roots. Comments about the quality of life--"Its a great place to bring up kids"--were made in many of the interviews. Cape George is a tight-knit community with modern recreational and cultural facilities. New inhabitants are quickly seduced into permanent residence.

Driving through Cape George one is struck by the obvious dependence of the economy on resources. Fish boats and guide-boats cluster in the harbor. Almost all the men observed were in the attire of mill worker, fishermen or logger. On the walls of the Blue Bistro the fifty year legacy of sport fishing and logging was portrayed in historical photographs and artifacts.

Leaving the town harbor and driving west, one is quickly sealed from the influence of the sea. For a short while tourist homes dot large lakes and campgrounds abound. This quickly gives way to a terrain in which huge swathes of trees have been laid waste or whole mountainsides; slash and tangle replace once verdant monsters. Whereas some farming--of an orcharding variety--could be found around Cape George, further west the terrain is more mountainous. There is scant evidence of human habitation until one reaches Haidlik, a small fishing community on the west coast. Here there are also great contrasts: in the harbour, gleaming tourist yachts tower over small, hardy fishing boats. On the day in question wind blew from the open Pacific in a steady gale. Temperatures were at least ten degrees colder than in Cape George, and as a dense chilling fog moved in from the ocean the warmth of what had been a hot day was quickly dissipated. One had a sense of how unforgiving nature can be, and how close the town's inhabitants were tied to it through their occupations. Haidlik is a compact, robust, somewhat

parochial and down-to-earth community: the children unprepossessing and disingenuous. These qualities extend to two communities further down the road; Todrow and Canforth. However, each town has its own sense of identity, and as is so often the case in rural B.C. this rivalry has been expressed in competition for the right to house government offices and schools. Haidlik is privileged to contain the district's only grade 8-12 school.

A high percentage of native peoples live in the coastal area. They are an important part of the cultural life in all communities--making up 7% of the total population. Although relationships between groups was considered cordial, in the words of the superintendent of schools, the native situation in education is "very sensitive".

Approximately 6% of the population from a variety of ethnic groups speak a language other than English as their first language. Only 44% of the people (compared to a provincial average of 55%) have attained high school graduation. One teacher claimed that people in Cape George, for example, tend to have a "mill-town mentality".

The Educational Context

Educational organizations are shaped by, and in turn effect the larger community. Members of the Stakeholder Evaluation Team insisted that any educational evaluation recognize this concept. Any comparative data was to be interpreted in the light of the socio-economic context. In addition, all members reacted strongly when the superintendent described the district as "average". Indeed there was much evidence to show that the district had many characteristics that gave it a unique identity.

Two examples will demonstrate Inlet's individuality. In Cape George, each year approximately 35 students travel on exchange to a sister city in Abishiri, Japan. In addition students or teachers from this district were involved in educational projects with six countries in Europe and Asia. Student accomplishments were also impressive. Students from Inlet placed first one year and second the next in a national science fair. One student's project consisted of adapting the dimples on a golf ball to wind-sails on boats. He has since patented the concept. Yet the district certainly has its challenges and weaknesses; it was the purpose of the district evaluation to identify these.

Leadership style and impact in the Inlet School District is also multifaceted. Senior district staff and teachers respect each other as progressive leaders, but see school administrators as traditional. One of the teacher representatives stated, for example, that all instructional leadership either comes from the teachers or district office. Over half the teaching staff was taking part in a summer workshop on cooperative learning that a director of instruction had organized.

Stakeholder Team members characterized the administration as "humanistic and progressive," and described the superintendent (Bob Ney) as "the hub of the district". Indeed, the teachers wanted him to sit in on all negotiation sessions, so the committee struck to deal with bargaining was reconstructed. Bob Ney is viewed as being "concerned with educational issues, provided they are done his way", and has "good working relationships" with all groups. He is given credit for the establishment of school parent advisory groups, and holds executive council meetings weekly, at which representatives from the IDTA, CUPE, (the Canadian Union of Public Employees), the Secretary-Treasurer, and Directors attend. The executive council serves a communications and management function, rather than a planning one, for as Sandra Johnson, president of the Inlet District Teachers' Association (IDTA) stated, "teachers have no input into goals, or district philosophy statements--especially on an ongoing basis". Teachers are desirous of input into the purposes of the district.

Attempts are being made to shift responsibility for instructional leadership to the school, and thus to the administrative group. One district administrator characterized it as an "old boys group, characterized by sluggishness;...(they are)...difficult to move". A progressive thrust for change is coming from the teachers, but principals are not viewed as instructional innovators. Prior to the evaluation, leadership within the Inlet District could have been termed moderately progressive. Relationships between interest groups seemed cordial, even though not all groups are moving in the same direction. Yet, there was ample room for improvement. The teachers' association still felt isolated from the planning process, and the administrative group moved at a different pace and only peripherally in the direction of the other groups. In addition, there were indications that trustees were beginning to become too involved in management activities; wanting to "play with trains", as one

interviewee called it. Into these dynamics were then placed the "more difficult, more hard-line" first round of negotiations under the union-management model created by Bill 20; the Sullivan Royal Commission of 1987-88, and the IPS project itself.

In late December, 1987, the researcher received a letter from the superintendent of the Inlet School District stating that "Our district is seriously intending to undertake the Ministry of Education developed Evaluation Project". In mid-January the researcher and Mr. Ney met for two hours at the Ministry, and reviewed aspects of the project: the questionnaires from the first pilot, the data elements contained within the district's Information Profile, and the Nechako report (see Appendices A2, A3, & A5). Tentative plans for the school board and the IDTA to review the revised questionnaires were made. This was to be done as soon as Bob Ney received the revised survey instruments. Mr. Ney stated he would contact the researcher upon completion of this task.

Over the next two months Mr. Ney reviewed possible evaluation processes with the school board (more than one model was considered), including the original and the new drafts of the Ministry's questionnaire. The IDTA reaction was as expected. They claimed that questionnaire wording was vague and indeterminate; and expressed some apprehension as to "how it might affect them"--i.e., how was the information to be interpreted and used? Bob Ney pressed on, however, quite confident that the teachers would ultimately go along with the survey. On March 18 a letter delineating responsibilities for both the Ministry and the school board was sent to the school district (see appendix C1). It was returned--signed--on April 8, 1988.

But what were the purposes of the project? In the Director of Instruction (Frank Ford)'s opinion it was "a local concern to improve the quality of our education ..Public education has become a bit of a whipping boy, and we wanted to find out if it was that bad..." Bob Ney's stated purpose was that:

I wanted a measuring stick in place in the district. I went about and convinced the Board and the teachers that we needed one, and got a process with integrity that ultimately worked....I really wanted some data from which to make decisions.

Others perceived that the management team and the school board wanted to improve public relations for the district. As one of the teacher

representatives on the Stakeholder Team stated, "the trustees thought it was a good idea. I'm sure the public relations value was in the back of their heads". Another stated,

In the first place I was quite suspicious, because it sounded like 'Oh, great, we're going to take in all this information, whitewash it, and tell everyone how wonderful we are...that it was a PR kind of toy'....Now I don't think that...even if that was the reason, its changed now...its an honest attempt to discover the weaknesses in our district.

Management's belief that an evaluation could be used to convince the public of the positive quality of education in the Inlet School District gave exactly the opposite message to the teaching staff. To them it implied what teachers weren't doing a good enough job, precisely *because* the public had to be convinced. This, combined with apprehension about potential misuse of survey information and the "addition of one more thing to their job" caused teachers to be less than enthusiastic about the evaluation project. Some teacher apprehension resulted from the fact that the questionnaires were developed outside the district as part of a province-wide Ministry evaluation project. For purposes of expediency Bob Ney decided to use the questionnaires as they were--rather than accede to teacher wishes to reword them. He did not want to engage in lengthy, semantic discussions resulting in little substantive change to the survey's intent. To alleviate some of the teachers' fears, and to "officially" bring the teachers on side, Mr. Ney arranged a meeting to take place on April 20--a luncheon meeting--at which time three representatives from the IDTA executive would meet with the researcher, Mr. Ney himself, and Frank Ford, the Director of Instruction.

In preliminary discussions to the meeting the researcher suggested a stakeholder collaborative team approach to data collection, interpretation, and report writing. Also, it was decided to do the work during the summer so that teachers would not feel the push of time and additional responsibility. A small stipend was to be paid to participants. The IDTA was also to have the right to choose its own representatives for the committee.

Mr. Ney structured the luncheon meeting in an attempt to accomplish two things: first, to allay fears and engender support for the project; and second, to gain commitment to a plan of action with specific timelines. He did this by clearly having his agenda laid out; by taking an informal, conversational

approach to a very formalized agenda; and by having the follow-up plan developed prior to the meeting. He also coached the researcher with respect to what he should say to the group.

In a surprisingly cordial and certainly non-confrontational meeting, both purposes apparently were achieved by the end of the hour. Teacher acceptance of the stakeholder plan (which included a trustee, a parent, and the director as well as teachers) seemed to mollify fears about survey interpretation and wording.

The tasks that were agreed to at that time and the timeline that was followed is shown in Appendix C2. Initial meetings of the stakeholder team were scheduled for June; however, difficulty with the timeline on the questionnaire (i.e., format time--the district had added two questions to the generic questionnaire--distribution/ collection time, and analysis time were all longer than anticipated) meant that initial and subsequent meetings of the group were not until July. The remainder of the proposed timeline was substantially followed.

During the afternoon following the luncheon meeting of April 20, and during preliminary discussions on July 4, the role of the researcher was discussed and clarified. He was to be a guide-facilitator; chairing the meetings with prior approval of the process from the superintendent. A plan of action was designed beforehand; it was the researcher's responsibility to supervise the process through to the end. In keeping with the concept of local control, however, decision making as to the nature of key indicators, the form of measurement, the interpretation of the data, and the final form of the book was to be left to the Stakeholder Team. It was also agreed that the researcher could make notes on the process for the purpose of his dissertation study, and would be "evaluating the evaluation".

Description and Critique of the Evaluation

Procedures and Processes

The procedures and processes followed in the Inlet District holistic evaluation differ significantly from those employed by districts involved in the previous pilot. Although the general strategy of surveying the community and using key indicators of performance related to the Ministry

values framework was retained, the manner by which each was accomplished and then converted into a final report was quite different.

A number of weaknesses in the IPS's methodology had been identified in the pilot. One of these--experienced in Inlet--was teacher resistance to the questionnaires. There were also difficulties with data interpretation, workload, the lack of substantive indicators for some goals or attributes, questions about the framework's accuracy, and difficulties in producing meaningful follow-up reports.

The four questions presented in Chapter III on the methodology of the IPS are answered as part of the meta-evaluation component of the study (Chapter VII). In this case report and in the remaining two cases, however, information and analysis relating to a specific district's procedures and processes will be presented under headings appropriate to their experience. In the Inlet case study they are:

1. Laying the Groundwork.
2. The Stakeholder Collaborative Approach.
3. The Four-Goal Six Attribute Framework.
4. Key Indicator Development.
5. Questionnaire/Survey Information.
6. Data Collection, Analysis, and Interpretation Procedures.
7. The Evaluation Report(s).

The information in each of these sections is drawn from content analysis of interviews, meeting records, observations by the researcher during visits to the district, and from numerous formal and informal conversations. Evaluating the processes and procedures from the point of view of improving the evaluation system will be done in Chapter VII.

1. Groundwork

a. Description. Much of the information relating to this topic has already been presented in the section entitled "Context and Background". It is clear that most of the groundwork--ie., in terms of meetings and discussions devoted to understanding of the ramifications of the evaluation project--were between the superintendent and the school board, and between the superintendent and the IDTA. Little effort was made to gain support from the administrator's group, or to introduce the project to the public of the district. Even the district

staff had limited knowledge of the project. The IPS evaluation project appeared to be the superintendent's initiative, and groundwork efforts were directed at those groups perceived to be possibly problematic, or vital for its support.

b. Critique. In the Inlet School District evaluation, most stakeholders felt the groundwork procedures were inadequate. Of particular concern was the lack of public awareness of the intents and nature of the evaluation. One interviewee stated that "better community involvement is needed"; another that it would have assisted the ultimate impact of the evaluation if someone had "let them know the process was going on: the public was not necessarily aware". A third stated that the evaluation could assist with the gain of public support for education, "but that would involve more than just putting out a questionnaire".

Teachers, administrators, the school board and district staff could have been better prepared. The IDTA president stated that it would have helped if the project was "more clearly explained to various stakeholders", and there was "a better understanding as to how it is to be used". Teacher suspicion of the process was still in evidence in the first July 5 meeting. Doubts about the survey intents and the procedures used in the collection of survey data were expressed. Frank Ford stated that it would have helped to inform "the in-school people more thoroughly". Administrators were not represented on the stakeholder committee, and appeared involved only in the distribution and collection of survey instruments to teachers and students. Bob Ney himself stated that if he had to do it again he would "the next time, try to get a larger group's commitment".

2. The Stakeholder Collaborative Approach

a. Description. At the April 20 meeting it was decided to formulate a panel of teachers and others--stakeholders acting as partners to district administrators--to design key indicators, collect and interpret data, and write the final evaluation report. The Team was to be composed of the director of instruction, the superintendent, and four teachers: one primary, one intermediate, one junior secondary, and one senior secondary. In follow-up discussions suggestions to broaden the committee were made by the researcher. Between April and June, the superintendent decided to broaden

the team to include a parent representative and a trustee. He issued invitations of two different kinds. For the teachers, representation was to be voluntary; i.e., the IDTA was to choose their own representatives to send to the committee. The same opportunity was given to the school board. However, the parent invited to take part was identified through informal conversation between the superintendent and trustees. No administrators, students, CUPE members, non-parents, or employers were asked to sit on the committee. The committee, hereafter referred to as the Stakeholder Team, was as follows:

Superintendent	Mr. Bob Ney
Director of Instruction	Mr. Frank Ford
President, IDTA	Ms. Sandra Page
Junior secondary teacher	Mr. Don Gordon
Vice-President, IDTA and primary teacher	Ms. Joan Holmes
Junior secondary teacher	Ms. Eva Moran
Trustee representative	Ms. Helma McLeod
Parent representative	Ms. Sarah Johanssen

To offset the fact that the teachers would be working during their vacation months, Mr. Ney offered to pay them (\$2500.00 was budgeted for teacher remuneration). The money was motivation enough for one teacher to take part in the project. Two others (the IDTA executive members) were there as "watchdogs"--to protect teacher interests. Otherwise the process would have been seen as "another example of trying to get teachers, to trip them up", stated the IDTA president. The fourth teacher, Eva Moran, stated that her motivation was to learn more about education in the district.

The parent representative, Sarah Johanssen, was both surprised and pleased to be asked to be on the committee. She was unaware of the reasons for her choice, other than her involvement in various school activities. Helma McLeod, the trustee, was bemused at her choice. She felt that it would have been more appropriate to have chosen a more knowledgeable trustee--after all she was new on the board. The words she chose to describe her role in the process were "insignificant, shy observer".

The Stakeholder Team was introduced to its task in early July 1988. Bob Ney began by explaining the background to the evaluation project. The researcher then gave an overview of the lessons learned from the first pilot. Teacher interest appeared to be aroused on three points: first, when it was made clear that the Stakeholder Team would write the report itself; second, when it was acknowledged that the uniqueness of the local district would be respected; and third, when the objective to find key indicators of performance for the social, human, and vocational goals of the Ministry framework was introduced. The latter was particularly surprising because a negative reaction to any kind of measurement was anticipated. However, when the project was explained as expanding measures of performance beyond simple test scores, teachers saw that the project as dignifying the teacher's role in meeting the needs of the whole child. Don Gordon expressed it this way: "I'm learning a lot more about education... this is helping me be less confrontational, less strident in my emphasis on intellectual development...I must use opportunities to deal with the other goals when they arise".

Stakeholder participation continued throughout the four meetings for key indicator collection, data collection and analysis, and final approval of the report.

b. Critique. There were a number of strengths associated with the stakeholder collaborative approach as implemented in the Inlet District Evaluation. A first is that it encourages individuals to divest themselves of traditional roles and become more system-oriented. Three of the four teachers on the committee felt that the superintendent's motives were altered by stakeholder desire to use the evaluation process for system improvement. Eva Moran stated:

There was a tendency at the beginning on behalf of Mr. Ney for this to be a whitewash kind of paper. But I don't believe he felt that way at the end... I think he saw it as everybody else did, as an instrument for change...I think he resisted that at first.

So too the superintendent and director saw the stakeholder process as a method to alter teacher resistance to the evaluation initiative.

A second strength revealed in the Inlet District Evaluation was the educative nature of the process (see Appendix F5). People learned from each other and about each other. For example, Eva stated that "...it was good to have

the boss there, he didn't know about the E.S.L. (English as a Second Language) problem". The trustee and the parent both remarked on how much they had learned through the process.

It must be emphasized that while much of what was learned was factual (i.e., about the nature of the school system in Inlet) there was also much learned about the motives, style of operation, and dedication of each of the members of the Stakeholder Team. Respect amongst the group grew. Bob Ney wished, for example, that he could tape record the discussions in the meetings and play them back to the public; and Sarah Johanssen stated that "they were the kind of teachers you would like to have your son or daughter work with".

As a result of the positive interactions within the Inlet Stakeholder Team resistance to the evaluation project--and concern, for example, on behalf of the teacher representatives to safeguard their interests--slowly evaporated. In this sense the stakeholder method controlled the potential for the evaluation information to be used for particular political agendas. In addition, getting data interpretations from a variety of different perspectives--ie., male, female, teacher, parent, district staff--helped gain agreement on what the "truth" of a particular piece of data or situation was. And finally, having a variety of perspectives allowed for a rich source of ideas and methods for the development and measurement of educational performance (29 suggestions for social development alone).

The stakeholder approach as practiced in the Inlet study also contained a number of weaknesses. It is impossible for individuals to simply put aside roles altogether. Joan Holmes put it this way:

Even if you have a committee that is taken from a hierarchy, the boss is still the boss, and there is not the kind of equality one would like to have...because you come in to it knowing that this person has more power than you. Even though we believe that what is said there won't be held against you...you have that feeling and take it with you.

And as much as revealing individual strengths, the stakeholder process can also reveal other qualities. Bob Ney, for example, was noticed as a "take charge kind of guy"--a little pushy and directive.

Absence of stakeholder representation from all groups affected by the evaluation in the Inlet district appeared to be a drawback. The absence of a high school teacher created data interpretation problems, particularly with student survey results. And, as Eva stated "possibly if the administrators had been represented as one of the members of the team, then they would feel a bigger commitment to the audit". On one occasion, when interpreting student attitudes toward administrators, decisions were made to excise the information from the final report because committee members were clearly discomfited in interpreting potentially damaging statistics. Other committee members suggested enlarging the team to include graduates, employers, and non-parents. Frank Ford and the IDTA vice-president also expressed concerns about lack of official representation from the west coast of the district. A final comment about the team: without strong people involved, i.e., without high levels of knowledge, skills and abilities in individual stakeholders, it is likely that the accomplishments would have been curtailed.

3. The Four-Goal, Six-Attribute Framework

a. Description. Acceptance and use of the values framework as defined by the Ministry of Education in the pilot project was never a contentious issue in Inlet. Time for discussion and alteration of the framework was provided at the July 5 meeting. Most committee members found the framework enlightening and meaningful. Ultimately it was agreed to go with the Ministry titles and definitions, with two main exceptions. Instead of using the term "human development" they used "individual development"--arguing that the latter term expresses the intent to develop qualities relating to an individual's physical and mental health and as such is distinct from social development, whereas the term "human development" encompasses both forms of growth. Also, instead of the attribute of "cost-effectiveness", the group chose the term "fiscal responsibility". This decision reflected the Stakeholder Team's belief that ascribing costs to measures of student performance is presently impossible.

b. Critique. The Ministry Framework of outcome values was strongly endorsed by the Inlet Stakeholder Team (see Appendix F1). One of the teachers in the Inlet District indicated her support for the framework, stating that "it is wonderful that we help children grow in more than one way". One apologist

however was Bob Ney, the superintendent. He stated his support in somewhat equivocal terms: "They are better than anything else we've got because we have nothing else". Terming them "jargonistic", he did state that "there would be no way we could have gotten sense out of the data" without them, and "they were fairly accurate but didn't give the total picture...". On the other hand, when pressed for alternatives, Mr. Ney had none forthcoming.

Another advantage--noted by the IDTA president, Sally Page--was that the values lend themselves to local interpretation. She stated that the framework elements-- i.e., values--are the same everywhere, "but are actualized differently district to district". She went on to state that "education has provincial implications, and I guess the province as a whole has to know what is available to students on a province-wide basis". The Four-goal and Six-attribute provincial mandate statement did not restrict local autonomy in the Inlet district.

One difficulty in coping with the framework was distinguishing between the meanings of social and human development. Although the former clearly focussed on student needs in relationship to society, many of the qualities that are a precursor to the development of social skills--ie., positive self-esteem, confidence, pursuit of quality--are the prerogative of human development. Finding discreet measures of each was difficult.

4. Key Indicator Development

a. Description. Much of the time spent in committee meetings was devoted to key indicator development. This was a task that blended a variety of different processes. They were:

1. Gaining acceptance of a definition of key indicators.
2. Brainstorming a list of potential indicators (through inductive analysis).
3. Defining the indicator in a fashion amenable to reporting; and
4. Winnowing, reducing, and gaining agreement on "key" indicators.

Most of two days was devoted to key indicator development. The researcher facilitated the development process, recording suggestions and moderating discussions. The definition of "indicator" used for the process derived from the literature review and the practical necessities of the project. It was accepted by the group with little discussion. The criteria were:

Key Indicators should:

1. be measurable and have a good range of scores.
2. substantially characterize one of the values, or goals of the Four-Goal Six-Attribute Framework.
3. add to the existing information pool.
4. be independent of each other; in other words, they should not measure the same quality.
5. be something we can alter through changing our practices; i.e., through positive action;
6. be in existence; i.e., the information should be available to us at this time.

Committee members were asked, during a brainstorming session, to suggest measures that were a function of their own intuitive judgment, based on their role in the system. Then, with reference to the criteria of a key indicator, the list of 29 was culled to nine--and decisions made to gather information to characterize them. Indicators were defined through discussion and implicit agreement. Some were discarded or left vague because of the difficulty of operationalizing a concept (for example, the health of multi-cultural relationships). Shorter time periods were spent on the other goals or attributes, but the same process was followed. At the end a list of key indicators for each goal or attribute was compiled.

b. Critique. Stakeholder Team members in the Inlet District rated the key indicators they designed as highly positive, but they did not feel that they were as accurate as positive (see Appendix F2). Part of the reason was not so much the key indicators themselves, but the problems experienced gathering the data to actualize them. The brainstorming process itself was accepted with little complaint; for example, Bob Ney stated "I don't know any other way it can be done".

The strengths of the process as implemented in the Inlet District was three fold: (1) teachers accepted the key indicator measurements, even in areas of education that were "immeasurable" in the first pilot; (2) "in-house" development allowed for a sense of local control of the project; and (3) ownership of the results was created amongst the Stakeholder Team, and within the district as well. Frank Ford expressed the latter strength when he stated that "it's an excellent process...it gave control to the district--albeit we

came up with the same indicators other districts had, we had selected them...we are more likely to act on these indicators, because they are our own choices". Stakeholder Team members in the Inlet District felt that the brainstorming process itself encouraged creativity and a sense of teamwork. As Sarah Johanssen said, "...one person would say something, and that would trigger another...and before we knew it we had all kinds of neat ideas. There was no feeling that if your idea wasn't used, that you weren't part of the process".

Weaknesses were also noted. Some found the process rushed. For others the concept of key indicators was hard to absorb. At other times, because the process for culling the list of key indicators was left to implicit consensus, there was, as Sam Gordon stated, the possibility that decision making was "swayed one way or the other by a few people in the group...I'm not sure whose opinion was the most influential". Others felt that the timeframe and the fact that they were dependent on information currently in place was a problem. Others felt more effort should have been made to develop indicators across the total spectrum of student ability. Most measures focussed primarily on achievements of above average students, rather than accomplishments of the population as a whole. Finally, it was suggested that the process did not result in a proper range of indicators (i.e., to express the many facets of a goal or attribute).

5. Questionnaire/Survey Process

a. Description. Earlier it was stated that the questionnaire used in the Inlet evaluation was a revised version of the one employed in the Ministry's first pilot study. The IDTA found the wording in the questions objectionable, and were not provided with an opportunity to alter them. This was the motivation for the formation of the stakeholder panel. Two questions were added by the superintendent to the pilot surveys. These related to support for school taxes and public perception of district leadership. All seven versions of the questionnaire were used (see Appendix C3 for details of the administration procedures).

Some problems with the surveys did arise, however. Teachers were not instructed to return the surveys to administrators in sealed envelopes, so fears as to confidentiality developed. Tracking graduates was a very difficult and time-consuming task (they had the lowest rate of return), and as a result a

rather unrepresentative sample was obtained. Turn around time for the scoring and interpretation by EMRG (Educational Measurement Research Group) was slow, and the format of analysis was very basic (i.e., frequency scores per question).

b. Critique. This was the least liked of all the procedures employed in the district evaluation in the Inlet district (see Appendix F3). Even though most people thought it was important to gain insights into the attitudes of the school community, they had many suggestions for improvement of the processes employed in questionnaire use.

Inlet Stakeholder members felt that a major strength of the process is the opportunity to gain input of a verifiable nature from various educational constituents. As the parent representative stated, "the questionnaire gives responsibility for the evaluation to the community," rather than just the Stakeholder Team. Secondly, even if the data are not statistically significant, surveys allow for identification of major discrepancies between groups, such as the consistently lower ratings (by at least 10%) scored by employers on almost all questions.

There were many drawbacks to the survey process employed in the Inlet study. First, the wording of the questionnaire was not acceptable to all groups potentially affected by the results; teachers were against it from the beginning, and although they acquiesced to use of the results, they never felt comfortable with the process. Second, sampling procedures and collection procedures were suspect. In addition, it was difficult to compare different constituent attitudes on the same concept because the questions were not the same on each survey. A fourth problem arose when the committee wished to see if there were age differences: no respondent profile was included.

A fifth problem experienced in the Inlet District's survey process arose upon return of the results. The questions had to be sorted and associated with the values framework; this took a lot of time and debate. And finally, because the questionnaire had preceded the development of key indicators, a number of issues that might have been properly measured through a perception check elided analysis.

6. Data Collection, Analysis, and Interpretation Procedures

a. Description. Data collection took place over a three week period in July. After key indicators were designed, each stakeholder was assigned collection responsibilities by the researcher/facilitator. Collection assignments were determined by the individual's willingness and access to data. There was an intent to spread responsibilities equally around the group. Part of the process was to redefine the key indicators to correspond to the quality of data, on the assumption that at the next meeting those indicators that were not correctly operationalized would be removed from the total list.

Individuals experienced varying degrees of difficulty in accomplishing their responsibilities. Most of the problems related to data access: almost none of the data elements were available in the form desired, or they had to be collected through contact with school personnel. In some cases the information was not available: eg., the key indicator dealing with multi-cultural participation in sports eluded measurement. The time frame for collection did not allow for adjustment if difficulties were faced; for example, gathering information on student citizenship from elementary report cards was more time-consuming than anticipated.

The data that was collected was analyzed, culled, and interpreted over a two day period. The process is described in Appendix C4. When the interpretation analysis session was complete the Stakeholder Team then chose two members--volunteers, actually--to take on the responsibility of report writing. These were Bob Ney and Eva Moran. Eva volunteered for two reasons: first, she was an English teacher and felt she could contribute productively to the report; and second, she would keep the superintendent honest--ensuring that decisions made in the stakeholder group were maintained throughout the report.

b. Critique. With the state of the data base found in the Inlet school district, and in keeping with the timeline established for the project, the main strength of the process used was that it did gather information--albeit information often flawed--that was acceptable to the stakeholder group. It also provided for the conversion of that information into a series of indicator measurements that reflected the value judgments of the stakeholder group as

to what was important to report about the quality of education in the Inlet School District.

Yet Inlet Stakeholder Team members found the process very time-consuming and somewhat inefficient. Information was difficult to come by as most individuals who possessed the knowledge had left schools for the summer. Eva Moran stated:

I didn't realize how much I had volunteered for...three or four days of phone calls...the information in the school system is inadequate; if the goal is to measure the four goals of education, then there is other more important information we have to collect.

In some instances in-school personnel resented having to gather the data. Others were not available. In terms of data interpretation, people found decisions as to significance difficult, as they were often faced with raw numbers reflecting a single year's performance. Color-coding appeared to assist in interpretation, but without standards of acceptable performance decisions did not sit comfortably with some committee members. Finding patterns amongst survey responses without common questions was also difficult. The timeline for collection did not allow for refinement or reformatting of indicators. And the absence of a high school teacher or administrator meant that significance of survey data from the high school population was hard to determine.

7. The Evaluation Report

a. Description. The Inlet evaluation report was written during the months of August and September, 1988. Responsibility for the conversion of data to prose, graphs, or tables was assumed by Eva Moran and Bob Ney. Guidance for the layout, length, and design of the report was provided by the researcher. Tables and graphs were designed by a contract graphic artist. A decision to use pictures as "demonstration indicators", to bring statistical information to life, was made by the stakeholder committee, in keeping with what they had observed in the Nechako report. Choice of pictures was made by an employee of the Ministry.

Mr. Ney and Eva spent two full days in Victoria roughing out the language and tables for the report. The rough pages were then given to the computer specialist who inputted them in typed prose, visuals, graphs, or tables. A

deliberate decision to present information in more than one way--ie., visually as well as in terms of the written word--led to a great deal of experimentation. As each section of the report was designed in rough, they were faxed to Inlet for approval and proof-reading. On two other occasions Mr. Ney returned to the Ministry to refine and improve the report. Presentation of information became a major issue in the project.

In late September a final rough draft of the Inlet Education Audit was sent to the district. The stakeholder committee was reconvened to proofread and ultimately approve the report for publication. On October 4 twenty advance copies of the report were delivered to Mr. Ney so the school board could go over them on their meeting of October 6. After showing the report to his wife, and asking trustees to do the same with friends, Mr. Ney also made the decision to produce 3000 more copies for public consumption. This was a departure from the expected, in that the report had been written as an internal planning document--for educators, trustees, and the management team. A shorter, more simple report had been planned for public consumption, but with this decision, that project was curtailed.

Applications of the Inlet District Evaluation

Once the evaluation report was complete, the researcher ceased to be a facilitator in the exercise and assumed the role of observer. In previous discussions the researcher and Bob Ney had discussed follow-up strategies to apply the evaluation's findings for decision-making, maintenance and gain of public and professional support, and accountability. However, no written, formal plan of follow-up had been agreed to.

In this section of the case report the events that took place in the six months following the report's publication will be described. A critique of the impact and effectiveness of these processes will then be conducted.

Description

Follow-up activities include newspaper reports, district-wide distribution of the report, public meetings, and a variety of other creative activities. Each will be discussed in chronological order.

The school board initially intended to go on its annual fall retreat and conduct an in-depth analysis of the evaluation's findings--leading to

development of district's goals for the coming year. However, the retreat was cancelled due to an illness to the board chairperson's mother. As an alternative, a special board meeting was scheduled between 4:00 to 7:00 on October 6. It was solely devoted to analysis of the evaluation report.

In attendance at the board meeting were six trustees, the superintendent, the secretary-treasurer, the special education coordinator, and two directors of instruction. Also in attendance were the researcher and an assistant observer. Mr. Ney conducted the meeting. He had prepared a comprehensive response sheet for each board member, and related documents (i.e., the district's philosophy). The plan was to systematically take the trustees through the book and design an action plan for improvement to the district.

Timing quickly became a problem. Some trustees became confused by the socio-economic information provided at the beginning of the booklet. One even became angry because the superintendent's explanation as to how the educational level of adults in the district--only 44% of the population had high school graduation, compared to 55% in the province--was interpreted as labelling the district as "all dummies". Although this confusion was quickly remedied, it reflected a problem that was experienced throughout the evening. Trustee understanding, knowledge, and perspective varied greatly. As a result, Mr. Ney was not able to move at the pace he anticipated. If the purpose of the report was to generate debate on educational issues, however, it certainly succeeded in the school board meeting.

The section of the book relating to the four goals of education was dealt with in a reasonably comprehensive fashion. However, as seven o'clock approached discussion of the attributes was rushed, as the superintendent was speaking at a Rotary meeting on the same topic. Thus much information did not get a proper airing; neither was there time to construct any kind of an overall action plan for improvement.

Many meaningful issues were identified and discussed. For example, trustees expressed dissatisfaction with student performance in reading and writing, and problem solving, based on 1988 reading assessment results. They identified employer perceptions of education as a definite concern; and indicated dissatisfaction with student perception of their own self-esteem, and of staff caring and school leadership. Debates ensued as to whether perceptions should be the focus of change, or whether district practices

should be altered. Awareness of lukewarm public support for increased taxes (asked as a questionnaire item) made one trustee state that "if we go into a referendum, boy, we had better do a good job of public relations". At one point it was suggested that a goal of 5% improvement in the public satisfaction rating should be established for the district. At the conclusion of the meeting, the superintendent challenged both the trustees and the management team to independently design action plans to improve areas of identified concern. Other formal follow-up activities included school visits by District Staff. All schools devoted time at a staff meeting for distribution and discussion of the report. In all instances the superintendent or an alternate district staff member visited the school. Each staff member was given a personal copy of the report. No structured response process was employed; staff were simply invited to send their reactions to the school board office. No formal follow-up to these meetings has taken place.

Some informal follow-up activities did take place. After conducting interviews on December 5, the researcher visited a parent advisory committee meeting held at one of two junior secondaries in Cape George. The main topic on the evening's agenda was the district evaluation.

Ten people were present at the meeting. After a short introduction by the principal, small groups were formed and individuals were provided with both a copy of the Inlet report, and a response form. Each group was asked to discuss outlined questions and make a report back to the larger group. A final report to the school board was compiled from a summary of the various submissions.

An interesting aspect of the parent group's report was the similarity of the problems they identified compared to those of the school board. They were markedly alike.

At least one other parent advisory committee devoted a meeting to a discussion of the Inlet evaluation. This took place at the senior secondary school prior to the December 5 meeting; indeed, it was cited as the reason that the turnout at the junior high meeting was so low.

The superintendent did four other things in terms of follow-up that were particularly significant. A first was to distribute the booklets to schools and the city hall for public distribution--"a box disappears about every three

weeks or so". He also took them to all-candidate meetings, and to the Rotary club for discussion.

Secondly, the superintendent provided class sets to each secondary English department and asked the teachers to structure a cooperative learning unit to look at the report from a student's viewpoint, and to write a paper commenting on it. After being graded the papers were turned into the superintendent.

A third follow-up activity involved the media. Bob Ney took the report to the editor of the newspaper, and shared the results with him. Expressing distrust of the examination results, the newspaper editor stated that on a test of common knowledge, he was sure that the students would score at least 15% lower than their parents. The result was a public wager. A 50 question test, designed by the editor, was given to four parent advisory groups and four randomly selected grade 12 classes. The results showed a 1.2% differential in average score. The newspaper editor, in a public editorial, admitted he was wrong in his judgement of student performance. Both he and the superintendent also agreed to "do something more like that, in order...to get people to start considering education".

A fourth overture was made to the business community. Mr. Ney sent copies of the Inlet report to all employers, giving them three evenings whereby they might come to a discussion group to "talk about what your concerns are about the graduates we provide". Only two affirmative responses were received, however, and so the meetings have not been held.

Critique

One strength exhibited by the Inlet evaluation follow-up was the ability of the district management team to use the report as an educative tool. Experience with the evaluation report showed that when meetings are structured to request identification of problems and suggestions for district improvement, they can be gleaned from analysis of the report. Such suggestions can point towards productive change within the district--witness the need for improvement of the ESL program, and for improvement of reading and writing skills. Clearly the evaluation report can serve as a common knowledge base for decision making, for many of the suggestions made at the school board meeting were also suggested by parents at the junior secondary parent advisory meeting.

A second educative use revealed by the Inlet District's follow-up strategy was revealed by the newspaper stories and editorials surrounding the challenge between the newspaper editor and the superintendent. Information in the report provides a basis for public dialogue on the quality of education, dialogue based on substantive information rather than pure intuition or emotion. And when the information itself is doubted (as were the exam results in the Inlet report) the existence of such information can as Bob Ney stated, "lead to other interesting studies"--witness the contest between parents and students, and the English assignments in secondary classrooms, asking for student input on the education system.

The educative function of the evaluation process lead Stakeholder Team members in the Inlet District to recognize the need to change people's impressions about education. One of the Inlet trustees commented that employer and non-parent perceptions of reading and writing skills are "possibly a legacy of their memory of school". She also stated that modern life does not support the practice of these skills; so it is incumbent on the board to "let the employers and non-parents know that priorities have changed". Whether the last statement is true or not is immaterial; the fact is that administration has a responsibility to direct action at the changing constituent perception as well as actually changing educational practice. The Inlet report clearly is a vehicle through which to improve accountability. It presents, in one place, an overall synopsis of the district's performance in education to both employees and the public. Parents have taken the report; parents have read the report; and parents have made suggestions for improvement to the system based on the report. The opportunity to do the same was also provided to the business community.

What were some of the weaknesses of the Inlet district's follow-up strategy? The major limitation was the superintendent's tendency to rely solely upon himself to initiate follow-up action. Some of the the stakeholder group did not question this approach; Eva, for example, stated that its "future use depends on Mr. Ney". Don Gordon stated that "I've heard that Mr. Ney has approached the business people, so maybe he, or the management, will take the responsibility to do something with it". Others, however, felt it limited follow-up effectiveness. Joan Holmes, for example, stated that the evaluation report was:

...never presented in such a way to give it the buildup that it could have....in our school someone came in not knowing very much about it at all, and noticed I was sitting in the group and turned around and gave it to me. It was done like, 'Here's the report, have a look at it, I don't know much about it', instead of 'let's have a look at it, point out a few things'.

Sandra Page the president of the IDTA, also expressed some frustration about not being included in the follow-up process:

...when we were doing the report we talked about the evaluation team playing a role in how the report would be used with the public, and we've not met since we looked at the original report....

She also stated that:

...whenever people have input and ownership of processes, then you increase their commitment and purpose...the professional staff in this district really want a good education system, this could point out strengths and weaknesses to work upon but they must have meaningful input.

She concluded by stating:

Goals cannot be a piece of paper that someone trots out when you're being evaluated, but are some things that are a living, breathing part of a school--that someone coming in almost should be able to state for you after a day in the school. This is only true if they are developed by the people involved.

These comments might go a long way to explain Bob Ney's greatest frustration: why teachers "did not show the same interest as members of the public" in the follow-up to the evaluation report.

A third limitation found in the Inlet District was imposed by circumstance. As this is the first attempt at a district evaluation, everyone was new to the whole process. The first evaluation is an event, not a stage in a systematic process. As a result all activities, all uses are creative, but comfort disturbing acts. Without a clearly defined plan for follow-up (such as followed for the evaluation process itself) that can, through formal commitment to it overcome existing habits, there is a tendency for individuals to fall back on old habits (witness Bob Ney's tendency to follow his traditional leadership style, and the acceptance of this by two members of the stakeholder team). This has the potential for alienating those who want a new approach (i.e., the IDTA executive) and curtailing the intent of the whole enterprise, which is the maintenance and gain of professional support for improvement.

A fourth limiting factor to follow-up in the Inlet District was the press of immediate events. The first round of negotiations between the school board and a newly unionized IDTA did not provide a climate of warmth and trust within which to engender spontaneous acts of cooperative planning. It also robbed inordinate amounts of time from educational leaders' educational agenda. In addition, the stress of change imposed by this circumstance and also emanating from the Royal Commission has not diminished turbulence in the system, although some aspects of anticipated change are very positive. It is difficult for teachers and others to respond to an evaluation report in any kind of sophisticated manner when so many new ideas are coming at them from many different directions. Without an organized response strategy, including time off for deliberation and reflection, teachers would not give the audit the attention it deserves. This was not provided by the district leadership; nor was it formally requested by the IDTA.

A final contributing factor limiting the follow-up strategy in the Inlet District was the report itself. Relying on individual interpretation for significance, it required guidance to make sense of it. No response form was attached or built into it. As a result a structured response process was required for effective results. This factor, combined with the poor groundwork laid for the process, contributed to a lack of spontaneous responses. Another possible explanation for the lack of spontaneous reaction from the public is possibly due--as Bob Ney stated--to their certitude that the education system in this district is in good hands. It may just well be that in doing the evaluation and in distributing the report, people were convinced that the administration knew what it was doing, and was in possession of all the evidence it needed to construct a plan for future improvement.

Summary

The Inlet district evaluation was a partial success. Many of the added procedures and processes appeared to enhance the model's effectiveness, particularly the stakeholder collaborative approach and the local development of key indicators. Questionnaire/survey processes remained problematic, as did the collection of "objective" data by which to judge the district's educational performance. In terms of achieving its anticipated goals the model performed marginally. It also was clear that this first attempt at

employment of the IPS in a school district to develop a learning process that is ongoing was little more than an event; habitual use is a different matter and an administrative challenge that the future poses.

CHAPTER V
CASE STUDY NUMBER TWO:

The South Coast School District Evaluation

South Coast School District is located on the south west coast of British Columbia, next to one of the province's major cities. Containing 7500 students, it covers approximately 172,000 hectares.

South Coast is considered to be a medium sized school district in B.C. Five main communities predominate. Most people live within the two suburban areas of Cochrane and Mountbatten. As one travels west the communities of Mariposa, Tunglat, Johnson Lake and Fort Randall are progressively more isolated and resource based. Fort Randall, at the extreme western boundary, is approximately two hours by road from the school board office in Cochrane.

The South Coast School District considers its problems and accomplishments distinct from those of other districts. Its socio-economic context, geography, and educational dynamics give a unique flavor to even the most general of issues. This case study seeks to capture the individuality of the South Coast District. A reader can then compare his own situation and determine the transferability of the study to his own venue.

Context and Background to the Evaluation:

South Coast's major centers--Cochrane and Mountbatten, on the eastern extremity of the district--contain the great majority of the district's population. Of the total population of 35,574 people, over 80% are within seven kilometers of the school board office in Cochrane. These two communities are bedroom communities for the province's capital city and are predominantly residential. The majority of inhabitants are employed in service industries, or in white collar work within the neighboring city. The number of local small businesses and services industries are growing in the area. A unique factor in the local economy is the military college located there.

Economic life in the main centres is in direct contrast to the circumstance in the District's smaller communities. They depend, as do most areas of rural

B.C., on the industries of logging, tourism, farming and fishing for their livelihood.

Driving through South Coast one is struck by the urban-rural differences that must be accommodated within a single district. The areas of Cochrane and Mountbatten are more closely linked in character and outlook to the neighboring city district than to the western communities. They also demonstrate the characteristics of modern urbanity: car lots, MacDonald's, modern subdivisions, and commuter traffic jams. Yet Mariposa (30 minutes drive) and Tungalat (45 minutes drive from the board office) are much more sedate. The populace follows occupations and enjoys lifestyles almost in opposition to those found in the eastern communities. Johnson Lake and Fort Randall are little more than logging camps on the open ocean, although the tourist season brings a constant influx of visitors. Throughout the district habitation is confined to a narrow coastal strip along the main highway. As a result the area does not give the impression of being particularly remote, as dotted along its length are farms, summer cabins, fishing lodges, and numerous campgrounds and tourist attractions. The great diversity in the district is reflected in the variety of ages, occupation, and education found amongst the school board trustees.

In the South Coast District most of the homes and businesses are modest. The average assessed value of residential property per pupil is \$40,000 less than the provincial average of \$169,000. The greatest concentration of property wealth is found in the eastern communities. Educational background of the general populace is slightly less than that of the province as a whole. Unemployment Insurance claimants and the number of lone parent families are slightly lower as well.

In terms of other socio-economic factors the district is quite uniform. Ethnically, the district is primarily English speaking. 1.8% of the population speaks a language other than English at home, well below the provincial average of 8.3%. Of the school population, 1.1% is Native-Indian compared to the provincial average of 2.3%. Native Indian education is not a major issue in the district, but the need to serve this population does create challenges for the educational staff.

Eighteen elementary schools are found in the district. There are two junior high schools, a senior secondary school in Cochrane, a grade 8-12

community school in Tunlat, and three alternative schools. The different character of communities within the district is exemplified by the schools.

There is some tension between the two secondary schools. The secondary school in Cochrane has a greater economy of scale and therefore has better flexibility of course offerings and timetabling options. Its students do extremely well in terms of earning local, provincial, and international scholarships. It is often referred to as the "lighthouse" school in the district. Its high profile tends to draw students from the Tunlat area. Yet it appears that the latter, a grade 8-12 school, seems perfectly capable of providing students with an education equal to that found within many schools in the province.

Leadership style and influence in the South Coast School District is characterized by a series of contradictions and cannot be uniformly described. The superintendent (Bill Wilson) describes the district's style as participatory and democratic. He has established four district committees to operationalize these traits. For example, the new Ministry primary program was steered through the formation of a working committee of teachers¹. The use of a stakeholder collaborative team during the district evaluation project demonstrated the same commitment. A strong district-wide parent association also is in place.

Despite these formal structures some others in the district feel the management style tends toward the more bureaucratic, traditional end of the scale. Some district staff and board members operate in an authoritative fashion. Who the main player is and what the issue on the table is often determined the administrative style used. Bill Wilson is a low profile leader who is more comfortable working behind the scenes than in the public eye; his influence on district management may not be enough to ensure that major decisions are in fact decided in a participatory way. Although relations between teachers, district administrators, school administrators and trustees appear cordial, one does not get a sense of unanimity and purpose within the district--partly because of its diversity, partly due to the problems of

¹ This is the first initiative that directly affects the schools as a result of the Sullivan Royal Commission. Dual entry to Kindergarten and a new "non-graded" primary program were implemented in September, 1989.

negotiations during the year of the study, and partly due to the inability to create a district ethos that embraces everyone within it.

This state of affairs may also be, in part, attributable to ongoing problems with regard to budget and public support for education. In the last two years there has been a great deal of public debate with regards to education funding. At the time of the evaluation, South Coast was in the midst of a minor taxpayer revolt. Inhabitants were finding the tax burden great. Assessments were lower than in the nearby city, but the same quality of education was expected. This condition was exacerbated by the fact that a third local district implemented a very successful public relations program that has enhanced its public image. The teacher representative on the stakeholder evaluation team alluded to this situation when he stated that "we have more computers than Tsawaan'en, but you'd never know it from the press."

These dynamics provided the impetus for consideration of the IPS project. District leaders were of the opinion that parents were knowledgeable and appreciative of the school system's successes, but felt that few non-parents understood the accomplishments of the educational system. The evaluation was envisaged as a mechanism to gather accurate information about "what people out there really think," and as a way to gain support for education through improved communications with the non-parent population. To quote the teacher representative once more, an evaluation is "more credible than straight public relations."

In late January the researcher received a phone call from the superintendent of the South Coast District requesting that the two meet to discuss participation of his district in the second round of pilots. The meeting was held in the school district board office in early February (for a full chronology of events, see Appendix D1). A trustee had attended a presentation on the results of the first pilot and had recommended the concept to the board and the superintendent.

During the inaugural meeting the superintendent was concerned about two main issues. The first was whether the district could retain control over the evaluation process, or whether it would be Ministry controlled; and second, what costs and disruptions to the district would occur. He concurred in a general sense with the stated purposes (planning, accountability, maintenance and gain of public support for education), but stressed its

applicability to the tax issue. Bill accepted assurances that the district would retain control of the evaluation: all decisions with regards to indicators, to choice of framework, to data collection and to the actual format and wording of the reports would be left up to the district stakeholder committee. He glanced at the survey instruments and appeared generally satisfied. He also was prepared to commit the district to the resources required in the suggested contractual arrangement with the Ministry, once approval from the school board for the project was forthcoming. This was achieved in its March meeting. Upon receiving the go-ahead the superintendent delegated the project to John Lee, the Director of Instruction. The researcher met with Mr. Lee in mid-April. The two were already known to one another because of collaboration in another Ministry project.

At the initial meeting Mr. Lee appeared somewhat harried. Although he was cooperative, and in favor of the project in principle, he was somewhat resentful of the additional workload. The duties associated with this project were laid on in addition to his normal responsibilities. He also exhibited a tendency to want direction from the researcher, as opposed to assuming a leadership role himself. There appeared to be three reasons for this. The first was the imposed responsibility. A second was his own interpretation of his role; he saw himself as the "district manager" of an process agreed upon by the Board and the Ministry--not as an equal partner in its development. The third was his own lack of understanding as to the true meaning and purpose of the project. All three reasons he expressed in his own words:

I was the district person, therefore I arranged the meetings...I was supposed to know where I was going but I really didn't...the first pilot didn't make much sense to me. This was just an add on job for me...My job was...to work with (the researcher); whatever he wanted us to do, we did what he asked us to do.

During the first meeting discussions were held about the format of the stakeholder collaborative committee, the questionnaire process, and the timeline for the project. The committee was to represent a variety of interest groups--teachers, parents, administrators, the school board and district staff. It was agreed that the revised questionnaires from the first pilot would be used, but they would be analyzed by the Stakeholder committee so as to gain support from the teacher and other groups. Yet substantive changes to the

questionnaire were to be discouraged so as to facilitate the project timeline. The survey was to be administered by the committee (see appendix D2 for a sample of committee responsibilities). In addition, the project would aim for completion at September-October, 1988, at which time an evaluation booklet would be produced. Tentative dates for committee meetings were established, extending throughout the summer months.

Role responsibilities of the researcher and the district project coordinator were discussed in a general sense. The former left the meeting with the belief that he was to be guide-facilitator, but that the Director would assist with the design of the project as it evolved. It was the researcher's responsibility to design the major components of the evaluation, but in keeping with the concept of local control, the plan was to encourage stakeholder team decision-making with regards to the choice of key indicators, the form of measurement, interpretation of the data, and the final evaluation report format. Stakeholder team members were also to cooperate in the researcher's plans for meta-evaluation of the project. It was agreed that the meetings would be run by the researcher, but with prior approval of the agenda from the Director of Instruction.

Description and Critique of the Evaluation Procedure and Processes:

The general format of the South Coast District holistic evaluation were similar to those employed in the Inlet School District. The two projects took place concurrently. The case study will follow a similar format.

1. Groundwork:

a. Description: Much of the information relating to this topic has already been presented above. Very little groundwork was done, outside of gaining permission from the Board of Trustees to take part in the pilot. No formal meetings with the teacher's association were held. School administrators were informed by letter as to their responsibilities concerning survey administration, but no formal meeting was held to give them an in-depth understanding of the project. The IPS project seemed to be of low priority, for reasons the superintendent himself indicated: it was "one of many initiatives coming down the pike from the Ministry--it may be very important to them, but to us it was just another of many new concepts." Also, because the district

was looking for ways to improve public support for education it was "an opportunity for the district to get some extra money and assistance from somewhere else," and the offer seemed a little too tempting to ignore. The superintendent later expressed regret at not involving himself personally in the project "but the responsibilities and pressures of negotiations superseded the priority of all else." For these reasons the project's profile was not as high as it might have been.

b. Critique: In the case of the South Coast evaluation there is ample evidence to show that groundwork was insufficient. Events reveal the low level of knowledge possessed by individuals as the project progressed. The key players in the ultimate use of the information--the superintendent, the director, trustees, and even Stakeholder Team members--were only slightly prepared to either conduct the evaluation or utilize its results.

Dan Roebottom, the teacher representative, commented on the lack of teacher preparation. He felt that there needed to be "more involvement at the school level, maybe one visit per school so people know (about the project) ...the teachers in our district don't understand what this is all about." He also stated that "...parents who have read the document don't even really understand what an educational profile is...it is an abstract concept to most people who aren't involved in it."

One possible reason for the inadequate groundwork was the lack of personal involvement by the superintendent. He himself stated that "he did not feel a sense of ownership" in the project, and that "a sense of ownership is key to utilizing the results." A second reason is attributable to the difficulties experienced by the Director, John Lee, in understanding the goals of the project. His admitted lack of vision with respect to its goals and potential purposes meant that neither the scope of preparation nor potential use was perceived. Groundwork efforts are required to begin the process of ownership amongst all who may be affected by the evaluation's results; therefore, an understanding as to the evaluation's potential is vitally necessary for its success.

A third reason for the poor preparation was attributable to the fact that the groundwork was required at the year's end, a time when every one is very busy. Certainly John Lee experienced this problem. A fourth possibility is that the natural communication processes in the district (both formal and

informal) are not as democratic and participatory as the superintendent might have perceived. The project was "delegated"; the choice and role of the Stakeholder Team determined more by administrative fiat than through shared decision making; and formal, rather than informal communications processes seemed to be the only mechanism by which information about the project was disseminated.

In conclusion it appears that the South Coast District's efforts with respect to groundwork reflected one, some, or all of the following: (1) a reluctance to share decision making; (2) a lack of understanding how such processes are developed; (3) an inability to develop participatory mechanisms, or (4) a lack of commitment to the project resulting in time constraints and limited vision as to its potential.

2. The Stakeholder Collaborative Approach:

a. Description: At the April 19 meeting it was decided to formulate a stakeholder group representing teachers, administrators, parents, district staff and trustees. The process of choice was left to the director of instruction. Further membership in the group was to be at the district's discretion. No additions were made.

Committee members were chosen in a variety of different ways. The teacher representative was selected by the local teacher's association, after a phone call to the local association office. Ken Franks, the administrator representative, was personally picked by the Director. The District parent's association--South Coast Parent Education Advisory Committee--chose its representative when requested to do so by the Director. Similarly, the trustees designated Jane Holden to represent them.

The full slate of committee members was as follows:

- | | |
|---------------------------------|------------------|
| 1. Director of Instruction | John Lee |
| 2. Teacher Representative | Dan Roebottom |
| 3. Administrator Representative | Ken Franks |
| 4. Trustee Representative | Jane Holden |
| 5. Parent Representative | Shirley Eastwood |

The Stakeholder Team met in early May to sort out its task. The agenda for the meeting was established jointly by the researcher and John Lee. Committee members did not come to the meeting with an understanding of the project. Reasons for the district taking part ranged from "being chosen to be part of a Ministry initiative", to the researcher "...had to complete his PhD (sic).". Thus, the introductory portion of the meeting was longer than originally anticipated.

Much of the time at the first meeting was spent debating the purposes of the project. Attitudes toward the goal of building public support for education varied amongst members of the stakeholder evaluation team. John Lee viewed the project as a way of "verifying" the good things being done in the district. He also gave the impression that few such weaknesses would or could be found. Shirley Eastwood agreed that the evaluation could communicate strengths, but also saw the need to single out weaknesses and develop a plan to overcome them. Jane Holden informed the group that the district was just in the process of publishing a series of goal statements for the coming year. She was concerned that the evaluation project might contradict the work already done, and give the public conflicting messages. However, the group generally agreed that the way to gain public support was to identify district strengths and weaknesses, and to provide the public with a plan to maintain strengths and to overcome problem areas. It was agreed that the evaluation could be a source of verification for the existing goals, and that it would provide credibility for increased resource needs. In a sense the public relations goal and the planning goal became rolled into one.

The Stakeholder Team in South Coast was able to discuss, but not decide how the evaluation would ultimately be used. The final decision was to be made by the school board. Jane Holden and John Lee in particular were reluctant to allow the committee to assume any decision making power that was the prerogative of the trustees.

The stakeholder group continued to meet throughout the duration of the project (May to September) and assisted with questionnaire administration, key indicator development, data collection and analysis, and report writing.

b. Critique: Use of the Stakeholder Collaborative approach in the South Coast evaluation revealed some similar patterns of strength and weakness. One strength was how it encouraged individuals to divest themselves of their

traditional roles; as Ken Franks said, "everybody on the team put personal pushes aside," and created a "group dynamic." This group dynamic seemed to contribute to a sense of ownership amongst the Stakeholder Team members. Shirley Eastwood thought the group was "committed to it, very committed to it," and John Lee commented that "there was discussion, and it was our--the committee's report--when it was finished."

South Coast Stakeholder Team members also found the process a learning experience. As Dan Roebottom said, "My interest was to get a feel for what the district is all about." All stakeholder representatives rated the process as highly educative (see Appendix F5).

A weakness of the Stakeholder Collaborative Approach in the South Coast School District was that it created role conflicts amongst some members of the team. John Lee, for example, agreed with other stakeholder members that the process should be part of a larger planning scheme; but at the same time he was unable to divorce his role responsibilities as public relations spokesperson for the district from his responsibilities as Stakeholder Team member. Because he was one of the two authors of the report, it tended to be more public relations oriented than planning oriented.

A second weakness of the Stakeholder Collaborative Team in South Coast was the failure to include the superintendent of schools. This weakened the profile of the project, deprived the group of vital information, and reduced the sense of district ownership. A likely consequence was that the committee seemed very reluctant to direct the project, being quite content to follow directions themselves--either from the School Board or the researcher. Two Stakeholder Team members (Jane Holden and Ken Franks) felt the group had been too compliant; they felt the questionnaire's imposition, the timeline and the imposed structures of the evaluation were too constricting. These comments were made during post hoc interviews, however, and the concerns were not apparent as the evaluation proceeded.

A third weakness associated with the stakeholder collaborative approach established in the South Coast District was the limited decision making power the Team possessed. Because the project had to progress through a hierarchy for important decision making, and because some members of the hierarchy were not aware of the dynamics of the evaluation process, decisions as to eventual follow-up procedures and mechanisms for implementation of those

procedures were given little importance and were inconsistent with the purposes of the project.

A fourth weakness was noticed in the selection of Stakeholder Team members. The process did not attract formal constituent group leaders; nor was there a process developed to ensure that those interest groups were kept informed as the evaluation progressed. As a result ownership of the evaluation product (the report) was confined to the members of the evaluation team, and it was not transferred to members of their respective constituencies.

A fifth possible area of improvement is in the number and variety of Stakeholder Team members. John Lee suggested that the group could have been expanded to include a CUPE representative, or possibly a student representative. Shirley Eastwood commented, however, that "the number of people involved was a fine number as long as they were representative of their group...more people would just muddy the water." The requirements of the IPS are comprehensive enough to suggest otherwise, however; a larger team would likely be preferable.

Finally, the South Coast Stakeholder Team demonstrated that the knowledge, commitment, and skills the Stakeholder individuals bring to the table determine, to a great degree, its potential success. The more understanding they have of the purposes of the project, the more likely their role will expand to meet the concomitant expectations. Certainly a major factor delimiting the project's successful completion was the time each member had to devote to the project. This was a complex function of energy level and the value each individual placed on other competing responsibilities.

3. The Four-Goal, Six-Attribute Framework:

a. Description.: Time for discussion and alteration of the Ministry's values framework was provided in an early Stakeholder Team meeting. The four goals and six attributes were accepted by the Stakeholder Team as appropriate beginning point for analysing the district's educational performance.

Little discussion took place with regard to the Four-goal Six-Attribute framework's appropriateness. One reason for this was that John Lee had been involved in another Ministry project that used the same framework for analysis; he spoke in its favour. A second reason was that the district did not

have a mission statement of its own that could be either substituted for the Ministry value framework, or that could engender debate of a contradictory or supportive nature.

The South Coast stakeholder group expressed some confusion with respect to the possibility of clearly distinguishing between human and social development, and the ability to measure these values. Ultimately the group accepted the Ministry definitions for these goals, changing only a title. One attribute was substantively changed, however. This was cost effectiveness. It was renamed fiscal responsibility. The concern with this term was likely caused by the high level of anxiety with respect to the tax situation in the district. The group chose the term fiscal responsibility because they agreed that while it was possible to get measures of the quality of financial management in the district, determinations of the relationship between costs and quality of performance in education was virtually impossible, and potentially inflammatory in the existing circumstance.

b. Critique: In the South Coast evaluation, the Ministry framework was heartily endorsed. This was likely attributable to two factors: (1) the district did not have a competing mission statement; and (2) the Director, John Lee, was cognizant of the framework, and strongly supportive of it. Little debate was required to gain their acceptance for the evaluation project.

John Lee expressed his support for the Four-Goal and Six-Attribute Framework when he commented, "Those are good goals...Where else can you go?" Don Roebottom agreed, stating that the "goals and attributes are a good framework to work within."

The committee saw the titles of "human development" as problematic, however. They used the term "individual development" instead of "human development", but accepted the existing definition. The group's problem with the term "human development" was occasioned by the apparent overlap with "social development"--and the title change diffused that difficulty for them. Concerns with the attribute of cost-effectiveness and its conversion to fiscal responsibility was the other main change.

4. Key Indicator Development:

a. Description. Key indicator development consumed most of the time spent in stakeholder team meetings. The group discussed characteristics of a good key indicator and a number of criteria were agreed upon. They were:

1. each indicator should clearly reflect the goal it is intended to represent.
2. each indicator should be independent as much as possible from other indicators.
3. each indicator should reflect what the public wants from a quality education system.
4. each indicator should be alterable through positive educational action, and the indicator should reflect these efforts;
5. each indicator should reflect issues of importance and or significance in characterizing educational quality;
6. each indicator should be an outcome measure of goal performance.

The second stage in the development of indicators was to brainstorm potential "outcome measures" for each goal. The rules guiding the brainstorming process were few, but important to adhere to. The first was that each person was to make suggestions that derived from their own intuition--that is, criteria that one employs regularly to make personal judgments about the quality of the system. Related to the first rule is the second: each person was to rely on those judgments that are a function of their own role in the system. Team members were not to anticipate others, or stifle suggestions out of fear that they reflected, for example, "only a parent's point of view". A third rule was that any and all suggestions would be accepted--no criticism or discussion regarding the legitimacy of statements which were generated during the session was to be allowed. People were to call out suggestions as fast as the researcher could write them on large flip charts. When the frequency of suggestions slowed significantly, a halt was called to this stage of the process.

Once the set of potential indicators was generated, the group examined and edited the statements to ensure clarity of definition and expression. Group judgments as to indicator quality were created in two stages. First, suggested indicators were examined according to the criteria established earlier. Second, the practicality of finding appropriate information to operationalize the concept was discussed.

Debate around the choice of indicators for social development demonstrates some of the issues related to the two stages of indicator judgment. For example, the "teen suicide rate" was suggested as a possible measure of human development, and "suspension rates" as a potential indicator of social development. The former was rejected, the latter accepted, according to the following arguments.

The suggestion of teen suicide rates was made because some stakeholders assumed that the number of students who take such drastic action indirectly reflect the system's ability or inability to enhance student emotional and psychological development. Others countered by saying that although this is true, one must question whether the suicide rate would (1) fluctuate enough from year to year so as to properly reflect whatever efforts the school district would make in the area of human development, particularly in a district this size; and (2) really be affected by positive educational action on behalf of the schools. The disagreement did not resolve itself through amicable discussion. As a result the researcher had to impose closure on the debate. When the Q-sort (see Appendix D3) was employed to prioritize indicators the teen suicide rate was not chosen as a significant indicator.

Debate on the issue of suspensions as a measure of social development also demonstrated some important issues surrounding the development of key indicators. Some Stakeholder Team members hypothesized that dismissals and suspensions are a result of students exhibiting either poor citizenship or a lack of respect for their school community. A high suspension rate therefore reflects deteriorating student respect; a low suspension rate mirrors improving conditions. In the end, the group agreed that suspension rates are an appropriate statistic to use as a key indicator of social development.

But agreement on suspensions was not achieved without vigorous debate. Ken Franks argued, for example, that if one uses this as a measure of system quality educators will simply react by trying to create no suspensions--possibly through lowering standards, or changing policy.

To overcome the problem of improper action to improve suspension rates, three suggestions were forthcoming from the group. First, it was suggested that this indicator be compared to the survey questions re respect and citizenship asked of students, parents and teachers. Suspension rates were deemed to be a quantitative measure of social development; satisfaction ratings

a qualitative measure. Together they provide a clearer picture as to the state of social development, and also help to minimize abuse of statistics. For example, if suspension rates drop, and the qualitative satisfaction rating of constituents also goes down, then two conclusions are possible: (1) respect and tolerance is not necessarily improving; and (2) whatever caused the reduction of suspension rates was not effective in improving social development.

Second, it was suggested that suspension statistics are only meaningful viewed in terms of trends over time. The long term provides a context in which to view statistics for any given year. Unavoidable variances in performance would soon be identified as a band of insignificant variance, and significant alternations in performance could be much more easily identified. With regard to suspension rates, these statistics were available for three years (whereas most other indicators were not), and the availability of the data helped characterize the indicator as a time trend, which was more meaningful.

Finally, another stakeholder suggested that the most effective suspension statistic would be one that did not encourage a self-fulfilling state of affairs. For example, a year of many expulsions would create an automatic improvement of the statistic in the coming year because most of those students would no longer be in the system. Therefore a statistic reflecting relationships amongst the captive population--i.e., school-monitored suspension rates--was deemed to be a more realistic measure of social development.

The two examples--teen suicide rates and suspension rates--illustrate the kind of judgment and decision making required of the group during the development and choice of "measurable" key indicators. At times the researcher had to be quite directive in bringing the sessions to closure. The Q-sort was helpful in this regard. The Q-sort was also used to give each individual equal input into the final assignment of priority. This allowed "key" indicators to be identified; and since publication of the final report required a limit to the number of indicators, prioritized ratings helped determine which indicators would be published and which would not.

On other occasions the group struggled with concepts that they could not put into measurable terms. In these cases a decision was made to include "demonstration" indicators; pictures or descriptive comment on the value being measured. Examples of these indicators included photography showing

the quality of student-teacher interaction; and lists of scholarship achievements.

Time constraints affected indicator development. Meetings had been originally scheduled for mid-July to complete all goals and attributes. The data was to be collected before the group met in August. Late scoring of the questionnaire and vacation plans disrupted the timeframe. Key indicators for intellectual development and the six attributes were not discussed until the late August meeting and were completed in a rushed fashion.

b. Critique. Support for the key indicator development process as it was employed in the South Coast evaluation was mixed. Two individuals felt quite positive about the process, and three felt completely ambivalent (see Appendix F2).

South Coast participants felt that the brainstorming component of the indicator development process was a positive one. Ken Franks stated, for example, that "I liked the brainstorming method to get the key indicators to start with." Jane Holden commented that "I like the idea of the brainstorming..." and Dan Roebottom said "...the brainstorming and the other processes to tie the indicators to your school district worked quite well."

There were a number of criticisms directed at the key indicator process by the South Coast Team. The first related to time constraints. Shirley Eastwood argued that "it was very slow...necessarily so, however, I think, if you want to do a good job." John Lee stated that key indicators "are useful," but "time-consuming." He also suggested that to streamline the process "maybe the Ministry should tell us which ones to choose". Time constraints--and the perception that the Team could not alter the overall timeline of the project--caused Ken Franks and Jane Holden to feel that the key indicator process was "too directed". "We didn't have enough leeway", Ken stated, "I really feel that a lot got glossed over...". Yet he did go on to say that if this problem was solved, and "you are looking at the method, then the method would have worked fine".

Two other areas for improvement of the key indicator process were noted by South Coast Stakeholder Team. Jane Holden felt that not enough attention was given to designing indicators of student performance that reflected the achievements of the whole range of students: "the list of indicators seems to target academic students only". Dan Roebottom had difficulty "trying to relate the information in the survey to the indicators...there seemed to be some kind

of incongruency there". The former comment may reflect the fact that no one on the Stakeholder Team, as an educator, represented the special needs population of the district. The latter comment is directed at the methodology of the evaluation project itself.

5. Questionnaire/Survey Information:

a. Description: A major task taken on by the Stakeholder Team was questionnaire distribution and collection. A decision to utilize the revised questionnaires from the first pilot had been made by the Director of Instruction and the researcher in their first meeting. The Stakeholder Evaluation Team assumed responsibility for distribution, collection, and interpretation of these surveys.

Much of the time in the Stakeholder Team's first meeting was devoted to laying out the procedures and timeline for the surveys. The questionnaires were to be self-administered. Sample sizes, collection procedures, and integrity of results were discussed and procedures designed accordingly (see Appendix D2 for details). It was agreed that the timeline did not allow for significant alteration of the questionnaires. Therefore, little time was devoted to the wording of the questionnaires, even though some concern as to the validity of the questions was expressed. However, the group did add two questions to the survey. One dealt with the tax issue and another the quality of district leadership.

Stakeholder members devised processes to distribute questionnaires through the schools to parents and students (elementary students were not to be surveyed). The district parent association took a leading role in this exercise. The Director was to assume responsibility for surveying graduates; Jane Holden, the trustee, took responsibility for the employer surveys.

Work began when the blank survey forms were received from the contractor. The process was delayed due to tardiness on behalf of the supplier. Responsibilities were verified in letters from the Director to the stakeholder team and to others in the district assisting with the process (i.e., principals and teachers). Administration of the survey took place in the latter part of May and the first half of June. Scoring and analysis of the survey results was carried out by a contractor. Analysis of the results was postponed until the August meeting of the Stakeholder Team.

b. Critique.; A post-mortem as to the success of the South Coast survey process was carried out upon its completion. Almost all sample groups had reached the required respondent size. Two exceptions were the graduates and the employers. Concern was also expressed that these two groups were somewhat non-representative in that only local graduates were available and only small business employers were surveyed.

A strength in the South Coast survey process was, as Jane Holden stated, that "it gave a real good insight into what the feelings (of the community) are". John Lee supported her but expressed the general viewpoint of the group when he argued that "the idea of using surveys is excellent; but the surveys themselves must be better. We could have got another 30-40% out of it."

Aspects of the questionnaire process that could be improved in the South Coast evaluation relate to groundwork, timing, questionnaire development, questionnaire administration, and linkage of the questionnaire process to the other procedures of the evaluation. June was a poor time for administration of the questionnaire--January or February was suggested as a better time. Stakeholder Team members resented not having enough input into questionnaire development: Shirley Eastwood stated, for example, that we were "not allowed to change a word on the questionnaire...we were told not to."

A second problem with the wording and structure of the questionnaires employed in the South Coast evaluation was the lack of parallelism between questions on the various constituent surveys. For example, Shirley Eastwood commented that "converting survey data was extremely difficult because of the survey that was used...so much of it had no application at all, so much had application to one group and not others." John Lee stated that this problem created "an attitude change from beginning to end...as we went along we saw that it wasn't going to be as useful as it could have been."

Third, although the South Coast Stakeholder Team was satisfied with the questionnaire administration process, it placed a large burden upon them. Representativeness of the survey results are suspect because of the corners cut to get the job completed in a reasonable time. The necessity to rely on voluntary labor severely limits the potential for a quality result.

A final weakness of the questionnaire process as employed in the South Coast evaluation has been alluded to already. In the case study section on key

indicator development it was noted that the group came to an awareness that the survey was in fact the source of "qualitative" indicators of performance, based on constituent perception. These were seen as distinct from quantitative indicators, and important for proper understanding of the goal being measured. These observations suggest that survey questions needed to be designed in the same manner that the key indicators themselves were designed.

The whole concept of taking an externally developed questionnaire and applying it to a particular district's holistic evaluation should be challenged. Ken Franks stated, for example, that a negative aspect of the evaluation process was "trying to standardize a survey form when districts vary so much, physically and economically;" while Shirley Eastwood described the evaluation as irrational--"because of the questionnaire."

6. Data Collection, Analysis, and Interpretation Procedures:

a. Description: Data collection for social and human development indicators took place during July and August. responsibilities were agreed upon in the final meeting before summer break. Collection assignments were guided by each person's willingness to work and ease of access to information. An attempt was made to distribute responsibilities equally amongst the group. Individuals were directed to try to gather information to actualize the chosen indicator. However, if some judgment was required on their part to alter the chosen indicator but retain its intent, they were free to do so.

Data with respect to intellectual development, vocational development and the attributes was collected following the August meetings of the stakeholder team. Only two weeks were available for this data collection stage. The project's timeline limited flexibility.

Stakeholder Team members commented on the difficulty of acquiring data during both timeframes. Few administrators were available for consultation during the months of July and August; and when they were they did not always react cooperatively. Some members of the Stakeholder Team went on vacation and had limited time for data collection. On many occasions the desired information was not found in a format that was useful.

Survey results and key indicator data relating to social and human development was interpreted and analysed by the stakeholder group during a

late August meeting. A significant amount of discussion revolved around the process of making and communicating judgments. With respect to the survey results, the group finally agreed that an 80% "strongly agree," or "agree" satisfaction rating, would be used as standard for delimiting a strength. However the percent level for designation as a weakness was left to the discretion of the individual doing the interpretation, to be verified during a round-table session later.

Each person was assigned two framework elements and large papers were placed around the room for pasting and collating data. Stakeholder team members were asked to affix data elements under the framework headings. Stakeholders cut and paste questions from the surveys, placing them under appropriate framework headings. Comparable questions from each constituent group were placed side by side. Stakeholder team members were asked to color code the information according to the scheme outlined in Appendix D5. Each individual then reviewed their judgments with the remainder of the group. Disagreements were aired and final decisions on coding were made.

District-designed indicators for which data was available were also analysed. If the information did not seem to do justice to the agreed upon concept, the indicator was rejected.

On the suspension issue a vigorous discussion ensued about the appropriate way to express the information. Should "suspension rates" be gross numbers of suspensions per year, numbers of student hours lost to suspensions per year, or percent of total possible student hours lost to suspensions per year? Should the unit of time lost be hours or days? Should all suspensions be lumped together, or should only certain categories of suspensions be used--ie., students suspended to the School Board versus school sponsored suspensions, in-school versus out-of-school suspensions? Should statistics be reported per school or for the district as a whole? What form should the graph take?

This discussion revealed an awareness that the form of the statistic and the visual representation of that statistic alters its meaning. Objectivity--ie., a singular representation of data that was correct, or fair--seemed almost impossible to achieve. In this particular instance truth was relative to the intent and perspective of the individual designing the information's format. As soon as one has to collate any raw data into a more generalized form, its

integrity is almost inevitably altered and its meaning open to new interpretation. In these circumstances a group compromise was required. The distinguishing line between information of fact and information of value seemed to be truly a negotiated line, irrefutable truth seemed hard to find.

At the end of the August meetings the group chose two members to assume the responsibility to gather the remaining data and to write the report. They were John Lee and Shirley Eastwood. They were to work with Ministry writers and designers to convert the information into an evaluation report. Meetings to begin this process were scheduled to begin in early September. The stakeholder group was to meet to approve and edit the rough draft of the report as soon as it was available.

It quickly became clear at the September meetings that information required to actualize a number of the key indicators was not going to be available. Stakeholder Team members and school administrators were too busy with school start-up procedures. Some key indicators were therefore lost in the production of the evaluation report.

b. Critique; In the South Coast District evaluation a pattern similar to that of the Inlet evaluation was noted. The process's main strength is that it succeeded in reducing huge amounts of information into a manageable format for the production of a final report. Its main weakness was that the quality of that information was questionable as to its integrity, and therefore significance.

How did the data collection, analysis, and interpretation procedures employed in the South Coast District succeed in converting masses of information into a report format? First, the data collection procedures provided the raw data to determine each indicator. Second, the analytical process provided criteria for examining the ultimate significance of the information. Finally, the interpretation process provided a mechanism through which each individual in the Stakeholder Team had input into the final decision as to how the data should be presented, and which data should be presented--thus, the report in its final format was endorsed by all².

² It should be noted that the endorsement precludes any endorsement of the report's objective validity or merit, due to the nature of the process.

A number of factors contributed to the poor quality of data in the South Coast evaluation. These factors were due to an interaction between the chosen implementation process and aspects of the strategic design. For example, timing problems were created through design flaws (i.e., scheduling the evaluation in the summer months); and implementation gaffes (allowing meetings to be rescheduled and data collection to take place at the wrong time). A second cause of poor data quality was the process for the questionnaire. A third source of weakness was the nature of the data found within the district: this was neither a design nor an implementation flaw, but a condition of circumstance attributable in part to the novelty of the project.

The first weakness--timing--created a number of problems. As John Lee stated, "you don't want to have to ask administrators and others to gather data during the last or first week of school". This was, of course, not the original intent, but the complication was as result of two reasons: (1) the excessive time taken by the contractor in preparing and scoring the questionnaires; and (2) Stakeholder Team summer vacations that caused the cancellation of two scheduled July meetings. This led to a need to gather almost all the data in the last week of August, during a time when all the Stakeholder Team members were themselves very busy with school opening procedures. A final problem connected with timing was that sufficient time to resolve data presentation concerns (typified by the debate on how to report suspension data) was not available.

In the South Coast district almost no existing data--extending over time--were readily accessible (an exception was suspension information). Most of the information needed for key indicator measurement had to be obtained from individuals in the schools. Much was not available in an useful form. In addition, some material that was collected and aggregated "disappeared" between two meetings, and the quality of the data available for report writing was lowered.

A final problem--and one that may be insolvable--relates to the conversion factor, or generalizability factor in information management. Taking data from its rawest stage into ever-increasing levels of generality, and presenting it in graphs and/or tables and/or prose, seems to alter its meaning. Can data integrity be maintained in this instance?

Most Stakeholder Team members in the South Coast District wished the data base to be improved. Jane Holden stated that a major weakness in the study was the need for "better, more accurate information", and that the evaluation project was a "basis to work from for improvement" in the establishment of such an improved data base.

7. The Evaluation Reports:

a. Description: Ten meetings took place over a three month period to bring the reports into production. The South Coast team members wanted to create two reports: the first, a public document (similar to one produced earlier in the first pilot) to be distributed to all homes; and the second an internal one along the lines of the Inlet report. The internal document was written first, during the months of September to November, 1988. Coordination of responsibilities amongst members of the Ministry team and painstaking attention to detail by the district representatives slowed the process down and created some production problems. Suggestions for improvement were forthcoming from the district representatives, but mainly in the area of detail--i.e., editing, picture selection, and appropriate prose. The full Stakeholder Team met during the last week in September to make a number of final suggestions to the general layout of the report.

Final decisions as to the number of copies of each report and to whom they would be distributed was made at a school board meeting in early October. Trustees were introduced to the rough draft of the internal document. Only two or three trustees appeared knowledgeable about the project at all, and it quickly became apparent that the issue was not how many documents to publish, but whether to publish one at all. After some debate that involved a great deal of clarification and information-giving, a decision to print 500 internal documents was made, with the distribution to be decided by the school board upon their delivery.

Although the researcher attended the meeting as an observer, the Director often relied upon him to explain various aspects of the project to the school board. Bili Nelson, the superintendent, later stated that he realized at that meeting that the project had not received enough of his own personal attention. This factor may have contributed to some of the confusion experienced by the trustees, and to the lack of direction with regard to the

report's publication.

The internal document was published in late November, 1988. The South Coast School Board made the additional decision to publish a newspaper insert for the public in early 1989. Both district representatives and Ministry staff produced that document between November and January: it was published in February, 1989.

b. Critique: In development of the South Coast Reports some strengths noted in the Inlet case study were enhanced, some were maintained, and some were lost. Some weaknesses were corrected. Few new weaknesses were discovered. There are two main reasons for this. The first is that different reports were produced for different audiences. The second is that mistakes learned in the production of the Inlet report were able to be corrected (to some degree) in the production of the South Coast booklet and newspaper insert.

Strengths that were enhanced were the attractiveness of format and ease of reading (but at some expense to the overall integrity of the data). Strengths maintained were: (1) the appeal to different learning styles; and (2) the function of being a good information summary. An additional strength was the fact that the report published a values statement for the district, where no mission statement had previously been agreed upon. Dan Roebottom commented for example that "the one strength that sticks in my mind is that it gives every interest group an opportunity to see the whole picture...rather than teachers being involved in their individual school or individual concerns...it gives them an opportunity to look at the whole district."

A third factor--deemed a strength by some, and a weakness by others--revolved around the decision to not build recommendations for improvement into both reports. The purpose was to encourage readers to use their own judgment and develop their own suggestions for improvement. Dan Roebottom, for example, agreed with this strategy: "I think its effective to the degree that everybody that reads it is going to put their own interpretation on the data." The reports still "tell you what's wrong...(they) certainly suggests some areas where we could improve," Dan Roebottom stated. In so doing he was in agreement with the majority of the other Team members. However, such a strategy depends for its success on the development of a process that encourages analysis of the report and elicits suggestions for improvement. This was not carried out in the South Coast district.

Ken Franks differed with the remainder of the South Coast Team when it came to judging the integrity of the information contained within the reports. He felt that "the report(s) did not express the evaluation of the committee," and as "far as capturing the essence...that's God-like...you aren't going to do that, even in 3000 pages." Jane Holden stated that they "came out terrific" in terms of integrity; others generally agreed with her.

The South Coast District's strategy of creating two reports went a long way toward solving the problem of multiple audience appeal. The product given to the public was shortened and written in a language that was more palatable to the public. This allowed the internal report to be structured more for educators, and seemed to contribute to its use. The timeline and the attention to detail exhibited by John Lee, meant, however, that the changes to the format of the document were not extensive--although it was much less cluttered than the Inlet production.

Applications of the Evaluation Information:

Description:

Once the evaluation report was complete, the researcher ceased to be a facilitator in the exercise and assumed the role of formal observer. A number of discussions with respect to follow-up strategies had been discussed in stakeholder team meetings, but decisions with regard to report use remained the prerogative of the district and particularly the school board.

In the words of the superintendent, the evaluation report was basically "tucked away on the shelf". He qualified this statement, however, by acknowledging that the report had contained information that was a catalyst to decisive action on part of the district. In one instance a trustee committee was struck to look at developing a strategy to improve relationships between the school district, and its non-parent and employer constituents. In a second instance Shirley Eastwood, the parent representative on the stakeholder team, was later employed to develop a public relations package for the district--utilizing, as a base, some of the indicator information contained within the evaluation report.

The reports were used primarily for accountability purposes. Each teacher in the district was given a copy, and the public version was delivered to each home as an insert in the local paper. Visitors to the school board office were

able to get copies of the public report in the waiting room. No formal process was developed to use the report to gather either reactions or suggestions from various district constituents for the purpose of district improvement; nor was the report discussed in a formal sense by the school board in a meeting for that purpose.

Critique:

What were the strengths and weaknesses of the South Coast Evaluation follow-up activities? One strength was the effort made to enhance accountability through the district. The production of a newspaper insert summarizing the evaluation report--and delivering copies of the internal document to teachers, the district parent organization, and one copy to each of each home-school organization--were positive steps. John Lee stated, however, that if he was to do it again he would just provide each staff with a report, rather than give each teacher a copy.

A major weakness of the South Coast District's follow-up strategy was the overwhelming neglect it received. Decision-making with respect to the use of the evaluation information lay with the school board, and despite the Stakeholder Team's suggestions to develop follow-up strategies in the domains of building public support for education and long range planning, none were formulated³. Contributing factors were (1) the senior management team's lack of vision as to the project's potential; (2) their reliance on a somewhat traditional administrative approach; and (3) the lack of involvement in the project on behalf of the superintendent. This created an ownership problem for both he and the school board members which was never surmounted.

Two other factors contributed to the inadequacy of follow-up strategies in the South Coast District. The first, strongly noted by the superintendent, was the work demands created for senior management by the new structure of local contract negotiations. Second, reactions to the evaluation report--so important to the Inlet District superintendent--were never really valued in the South Coast District. This lack of interest in getting a response to the

³ Some decisions were made after the fact, however, to deal with these issues. The point here is that no follow-up was planned in advance, so any decisions that were made were done in an ad hoc fashion.

report--represented by John Lee's desire to minimize any controversial issues, and the superintendent's lack of interest--is representative of a traditional approach to administration. Rather than seek out information, and share analysis of the results in a participatory way, the district leaders chose to "tell it", in a one-way, "down the hierarchy" fashion.

Summary:

In light of the circumstances described in this case study, the South Coast evaluation can be described as a "qualified failure". The word failure is used for two reasons. First, there were many methodological factors that could and should have been more carefully implemented and that would have increased its impact significantly. A second reason is the simple fact that, in a formal sense at least, application of the evaluation's information was left to happenstance. This was in part due to what Hodgkinson (1978) calls the pathology of superficiality: senior leaders created the semblance of treating the evaluation seriously, but in reality treated the results in a superficial way. The failure is qualified, however, also for two reasons. First, the project was completed; accountability was certainly enhanced, and some informal spin-offs from the evaluation did affect the district's operation. Second, evaluation of the project clearly reveals a potential that was not realized, but that could have been achieved with minor modifications in attitude and approach. As a pilot study, it was marginally successful operation.

CHAPTER VI

CASE STUDY NUMBER THREE:

The Interior School District Evaluation

The Interior School District is located in the central part of British Columbia, about 500 kilometres from Vancouver, B.C.. It spans approximately 5000 square kilometers, has a populace of 90,000 and a school enrolment of 15,900. It is considered to be a medium to large district in British Columbia. Most of the employment in the area is associated with small business and tourism. The major community (Appleton) is the focal point of economic and social activity. A major highway links it with the four outlying population centers, which are little more than satellite communities--none is more than 25 kilometres from the city center. It is little more than an hours drive from the northern district boundary to the southern extremity.

The Interior school district is in the midst of going through a major transition: from being somewhat of a traditional, inward-looking educational organization to more of an innovative, community-centered one. The evaluation project was enthusiastically endorsed as a process that would enhance this general thrust, a thrust that has, according to authorities in the district, made some strides toward improving relationships within the educational community. Central to the new administrative style is its slogan, "Together We Learn". In the opinion of the district leadership, it was a concept consistent with the intent of the district evaluation process.

The evaluation project in the Interior School District was initiated in July, 1988 and completed (in terms of the Ministry of Education's involvement) in March, 1989. The general strategy followed in this evaluation was similar to that of the previous two case studies. Some major tactical changes were attempted, however. For example, the decision was made to alter the obviously flawed questionnaire process by engaging a professional research company to conduct a public poll on education. Conceptually, however, the events and analysis of those events can be presented under the same conceptual framework found in the two previous case studies. This will facilitate the formation of patterns and insights amongst case study experiences.

The Context and Background to the Evaluation

The main center in the Interior school district--Appleton--has a population of approximately 65,000. It is a modern, recreation-oriented community that has an urban/rural blend; the latter derives from the orcharding that takes place in the surrounding countryside. Appleton is located along one shore of a large Lake, and as such has a large appeal to summer tourists. The area's climate is dry all year round, and hot in the summer and cold in the winter; it allows for a variety of recreational pursuits. Appleton's satellite communities--Glendale and Orchard to the north; Vineyard and Fruitville to the south--all benefit from the agricultural and tourist industries.

Economically speaking, Interior could not be characterized as a wealthy community. Most income is generated through service rather than industry; average incomes are approximately \$6,000 below the provincial average of \$26,180¹. Unemployment rates are higher than the provincial average, probably due to the large amount of seasonal employment. Residential property values are only slightly below the provincial average however.

The social fabric of the district is also distinct. Although it has about the same ratio of students to total population as does the rest of B.C., the proportion of senior citizens is much higher (this helps explain the low per-capita income but average property value). So is the proportion of single parent families. Ethnically speaking, the district is quite homogeneous. It has a small native population, and English predominates as the first language in the home. The largest ethnic minority groups found in the district are German, Ukrainian, and French in origin.

To many British Columbians, Interior is a desirable place in which to live. Driving through the district one can understand why--the climate and environment are clean and hospitable; outdoor activities thrive; and there is the distinct sense of community identity, probably because many families have lived in the area for many generations. The area always seems to be bustling--partly because many of the main centers are on a major

¹ It is interesting to note that the Interior School District is the single largest employer in the area. It employs 1400 people with almost 900 of those being teachers.

transportation route, and partly because business is growing in the local area. There is evidence of progressive entrepreneurial growth--throughout the researcher's prolonged association with the district, growth continued apace. It is clearly an area that thrives on service to locals and to tourists.

The satellite communities to Appleton tend to be miniature copies of the main center. They all have the same economic base, and tend to grow at a commensurate rate. One hesitates to call them suburbs because they tend to provide services to the tourists and orchardists who are at a reasonable remove from Appleton, but they are suburbs in that a significant number of people working in Appleton commute from outside. One does not get the sense that the district is comprised of disparate centers of local interest--on the contrary, it seems quite homogeneous.

The Educational Context

The Interior school district has nine secondary schools, 35 elementary schools, and one alternative school. While all of the satellite communities have elementary schools, only Vineyard and Glendale are served by local secondary schools. In such a large district the district office is a large entity quite removed from the schools, and the term "loosely-coupled organization" appears an apt description of the district. In many ways Interior is a composite of sub-organizations more than it is a single entity. This is reflected in the dynamics of district administration. Senior administrators are constantly called on to deal with organizational leaders who themselves are capable and committed to their own subsystem's goals.

Earlier it was stated that the Interior School District is going through a transitional stage in its development. Two years previously the district and district management had been described as traditional, bureaucratic, and almost anti-education. Class-sizes, teacher salaries, and attitudes to teachers were extremely poor. Morale was at rock-bottom. To overcome this, the district evaluated its internal management structure and the decision was made to hire a new chief executive officer. He was chosen for his vision and his enthusiastic and energetic attitude to education. Since his appointment a number of new initiatives have begun, all directed at improving those aspects that the evaluation revealed as contributing to the district's malaise.

There were two major strategies employed to improve the district's leadership style. The first was to develop a structure of participatory management, including the development of a district mission statement and yearly goals. These structures were developed within the first year. A second was to improve two-way communications, not only internally, but also with the public. A cornerstone of the latter was to create a district annual report that would highlight those aspects of education that stressed the progressive, people-oriented thrust of the district's programs. A "district video", profiling teachers, students and staff was produced to carry the message to employees and populace alike.

It was into this context the district evaluation project was injected. In July of 1988 the senior management team in the district contacted the Ministry of Education and requested information about the Nechako report model, hoping to improve their annual reporting procedures. In the process of discussions a decision was made to work on a cooperative evaluation project that would (1) assist in the planning process, and (2) utilize both a district video and a booklet as two versions of a district annual report. Details as to the working arrangements were ironed out in the months of August and early September. Contractual arrangements were finalised in mid-August. The evaluation was to take place between September and December, 1988, with the final reports completed by early March.

Throughout the duration of the evaluation the new dynamics of contract bargaining came into play. They gave the district management team its first real test as to how effective its previous measures had been; as one teacher put it, "the bloom is off the rose". Questions were asked as to whether the new management style was "lots of words, but no action". The Central Interior District Teachers' Association (CIDTA) exercised its militancy, and throughout the latter stages of the evaluation project orchestrated a work-to-rule campaign and conducted a strike. An additional complicating factor was the Royal Commission Implementation Plan, which introduced a number of immediate changes to the B.C. education system. The ability of the new administrative style was (and still is) being severely put to test.

Yet through all of these disruptions the IPS project was brought to conclusion. The disruptions both strengthened and weakened its potential for enhancing leadership in the district. In some ways the context in which the

IPS was used was the context for which it was designed; much of the situation described above reflects many of the qualities that gave rise to its generation back in 1985².

Description and Critique of the Procedures and Processes
in the Interior School District Evaluation

The general format of the Interior District evaluation was similar to that employed in the other case studies. However, as the evaluation took place during and after the others, some evolutionary changes were incorporated into its design. This particular case provided the opportunity to try to implement some of the suggestions for improvement mentioned by Stakeholder Team members in the other school districts.

1. Groundwork

a. Description. From the beginning the IPS project was a high priority within the district (see Appendix E1 for the chronology of events). Both the superintendent and the District's communications director shared responsibility for its completion. Both were in regular communication with the researcher throughout the two months prior to the first stakeholder Team meeting. Jason Stone (the superintendent) took great care to understand the project's purposes, and to ensure that it was consistent with the style and directions envisaged for district leadership. For example, he stated in his interview: "...some of the questions presuppose that we hadn't considered this before...This is one of the thrusts that we (had)." He went on to say that:

I found it complementary with the planning process, that we had already initiated here with the statement of mission, yearly goals, etc. and the feedback loop that this would help us provide. So it was something that would allow the ministry to meet its needs, allow us to

² At the time of writing (eight months after the evaluation) negative feelings between teachers and the senior management due to the highly emotional contract negotiations have not improved. In this sense the IPS did not succeed in reducing district conflict; however, there is the possibility that its use has minimized the extent of the current disaffection between district leadership and the teachers. The district leadership continues to use the evaluation information and "key indicators" as a fundamental part of district administration (Damon Korchuk, personal communication, October 3, 1989).

meet our annual report needs, and dovetail with the planning process.

A first step in his groundwork activities was to take the project to the School Board--not so much so as to request permission to participate as to inform them that he would like a trustee to participate in the district Stakeholder Team. In addition all the major interest groups in the district--students, CIDTA, parents, administrators, Canadian Union of Public Employees (CUPE), and district staff--were notified about the project and asked to send representatives to the first Stakeholder Team meeting. The intent was to introduce the concept to them, and ask them to carry the appropriate messages to their constituents.

The researcher and the director carefully constructed an information package and presentation format to explain the nature of the project. A full day was put aside for presentation and discussion of appropriate concerns. Topics included (1) the necessity for team leadership in the district; (2) goals of the project--from both the Ministry and District point of view; (3) the nature of the evaluation process, including the use of key indicators of performance (survey and quantitative) and the use of a values framework as the basis of evaluation; and (4) the potential role and tasks of the district Stakeholder Team. Sample evaluation reports from other districts were handed out along with other explanatory material.

In this same meeting it was decided that the various constituent representatives would disseminate information to their respective publics. In addition the district assumed the responsibility to publish an article outlining significant details of the project in the main Appleton newspaper. The article accurately expressed the purpose of the project, the steps taken to characterize the district evaluation, and the timeframe of the project. Other groundwork efforts (or what might be better termed, "motivational techniques") were employed by district staff throughout the duration of the project. They made themselves accessible, provided resources, and prepared staff for the inevitable disruptions created as the project progressed. Jason Stone, for example, concentrated on developing a "positive working relationship among the people and a positive feeling about being there." Things like "a phone call," "...(sending).. a book to a person part way through, who was interested in a particular idea," "the lunches we had, talking to people

about where they are going, what they are doing," were all attempts to create an "understanding as to what everybody's agenda was". Damon Korchuk (the Director) also was sensitive to that and "was sending the letters and the follow-up and the thank you and the phone call." In Jason's mind the administration of the project had to "really work out the 'together we learn' motto."

b. Critique. Groundwork efforts in the Interior School District were moderately successful. Where they were particularly effective was in creating a good understanding within the Stakeholder Team as to the project's purpose and their role in it, and in engendering a positive attitude to the evaluation. Less effective were (1) the follow-up efforts on behalf of the various constituent group leaders, and (2) the attempts to create a community-wide understanding and commitment to the exercise.

A key element in the groundwork success was the efforts of the superintendent to establish a high profile for the project. He stated, for example, that:

I think if the CEO (Chief Executive Officer) isn't involved in something like this it sends a message...the message I wanted to portray to all participants was that this was an exceedingly important activity, and I would therefore commit myself to being there and being involved.

His efforts in morale building supported this intent. Is it pure coincidence that all Team members described their role as either important or somewhat important, and seven of the nine chose the term "positive" as a word that described their role well?³ Bill Blair, the school administrator's representative, supported this contention when he stated that:

...the members of the committee expressed, just through actions and commitment a definitive partnership in developing something that was worthwhile and positive for our entire district.

³ A second possible interpretation can be placed on the efforts of district staff to "motivate" Stakeholder Team members. Were there ulterior motives? One might question to what extent the phone calls and lunches were truly sincere efforts at creating ownership, or were subtle manipulations directed at seducing stakeholders to view the project in a positive light. Although one was left with the feeling that they were sincere, the possibility of careerist manipulation does exist.

A second possible contributing factor to the groundwork success was the clear link between the needs of the district and the purposes of the project. This link was enhanced in that the participatory "shared" leadership style of the district leaders was complementary to that employed by the general structure of the district evaluation model. The Director, Damon Korchuk, emphasized this link when he stated that:

We had heard that there was a process for evaluating the district that used a broad-based committee and that held a certain level of integrity...the key to this whole thing is involvement...it really formed a partnership, it was a team of people, and we were cooperating...I really like that.

Bill Blair gave further evidence of this link when he stated that "there was ownership for the development of this thing as it related directly as to what we are attempting to do in the district, and that ownership helped bring about a very positive attitude."

A third possible contributing factor to groundwork success was the careful effort put into preparing the introductory meeting to the project. Louise Loewen, the student on the committee, commented that "the groundwork in this case was generally well done. Everyone who came here ...(for)...the first meeting...were told what was expected." They "were fine." Norma Rivette, the CUPE representative, stated "I don't think we could have done any better than we did." On the other hand, Sonja Kocan, the parent representative commented that "I really felt disadvantaged the first couple of times...half the stuff they were talking about was eduspeak, or educationese, or something..." Groundwork efforts may have been appropriate for committee members who live in the system each day, but inappropriate for those somewhat distant from the schools.

Groundwork strategies were not as successful in transfer to various interest groups or to the community at large. Damon Korchuk stated that "not all the players in our system knew what was going on." Evidence existed to show that Stakeholder members had attempted to inform their constituent groups, but were not as successful as they would have liked. Norma Rivette stated for example that individuals who had not attended union meetings "weren't informed." Colin Scott, the trustee, did not feel the trustees were as well-prepared as they might have been:

...the very issue of evaluation should have debated in an in-camera meeting or something, or a retreat; so it should come out of a thorough think-tank session...I don't think the board was thoroughly committed to the process.

James Oord, the CIDTA representative stated that he wanted to improve "the forewarning of school staffs in my particular case". He also wanted to work out a better way of "getting input from our constituent groups." And although the community newspaper article was available for all to read, the single story had limited impact.

2. The Stakeholder Collaborative Approach

a. Description. The Interior District Stakeholder Team was formulated in the first month of the project. The complete list of Stakeholder Team members is shown below:

- | | |
|--|-------------------------------|
| 1. Chief Executive Officer | Jason Stone |
| 2. Director of Communications | Damon Korchuk |
| 3. Parent Representative | Sonja Kocan |
| 4. District Administrators' Representative | Bill Blair |
| 5. CIDTA Representative | James Oord |
| 6. Student Representatives | Louise Loewen
Chuck Morgan |
| 7. CUPE Representative | Norma Rivette |
| 8. Interior School Board Representative | Colin Scott |
| 9. Teacher Representative | Sid Pollack |

Selection methods for the Stakeholder Team were partially described earlier. Each formal sub-organization was asked to send a senior representative: in most cases, the president of the organization attended; the exception was the school board representative, who volunteered. Student Representatives were chosen by the Teachers' Association, and did not come as elected delegates from the student population. Sid Pollack, the second teacher on the committee, was asked to participate because he had the responsibility to produce the District annual report video.

The Interior Stakeholder Team met on six occasions over five months. The Team's role was to identify key indicators, collect information on those key

indicators, and convert that information into an appropriate representation of the key indicator's nature and significance. A related responsibility was to decide the values framework that would guide key indicator selection. A third responsibility was to assist in the development of a public survey; and a fourth was to help write and edit the evaluation report. Sid Pollack's responsibility was to develop a district video that would capture the evaluation results and that could be used to engender public discussion about the quality of education in the Interior School District. Although it was clear that the follow-up strategy (and it was discussed on occasion) was to be determined by District Management, the key players in district decision making (i.e., the superintendent, the trustee representative, and the Director of Communications) were prepared to accept direction from the committee in terms of report production and use.

Working relationships within the Stakeholder Team improved as the project developed. Bill Blair, the Interior District Administrators' representative put it this way: "The overall leadership shown by the many committee members...grew from somewhat hesitant to being very productive and sincere." Engendering discussion on issues was never a difficulty with this group. Keeping the meetings on track was taxing at times as many individuals were opinionated and wished to contribute on each issue. At times the student and parent representatives felt the discussions were lengthy and turgid. Without the Q-sort (for decision-making purposes), and without structured timelines and criteria to focus the discussion, the meetings could easily have gone more off track than they did.

A variety of procedures were followed to keep people involved and motivated to work. Strategies included: brainstorming; structuring small group sessions with reports back to the main group; ensuring that each individual was given the opportunity to give their opinion on important issues; ensuring that at times each person took the role of spokesperson for a small group; and assigning clearly defined tasks within a short time period. In almost all instances these tasks were evaluative in nature--they required the

stakeholder to use his judgment, based on the application of criteria to specific situations⁴, to make decisions as to priority rankings of alternatives.

b. Critique. Stakeholder Team members generally found the Team meetings to be productive and stimulating. There was one notable exception, however: Chuck Morgan, one of the student representatives seemed overwhelmed by the process and withdrew from the Team after the third meeting. He claimed pressures from outside activities were responsible for his decision. This may very well be the case; however, it is also true that the meetings were, at times, turgid, wordy, and suffused in "educationese". This made it difficult for the non-educators (i.e., the parent, students, and the CUPE representative). Even educators found the meetings "taxing"; Bill Blair stated for example that "I left each meeting feeling worn to a frazzle" But, he also stated, "It was a good feeling, always a good feeling".

The Stakeholder Collaborative interactions were such that the words most predominantly chosen to describe the group were a "positive, involved active-cooperating partnership" (see Appendix F6). Louise (a student) commented that "everybody managed to pull off two goals at once...kept the integrity of it, but at the same time had what they wanted from their viewpoint." The trustee, administrators' representative, and CUPE representative commented on how they were able to give productive input into final decision making. Sonja Kocan, the parent representative stated that "I liked how all the ideas came together, and how everyone helped decide...it was important." This is significant in that she felt lost during the first two meetings discussions, and described her role in the process as that of a "shy, cooperating participant."

Stakeholder Team members were not as positive about the process of representation, nor of the opportunity for them to communicate with their constituent groups and truly represent their viewpoint. Both district staff members questioned the ability of the students to adequately represent their constituency. They implied that the exercise may have been beyond their maturity. Others mentioned that representation from "retirees", "the Chamber of Commerce", or "employers" would have been valuable. Colin Scott, the

⁴ An example was when the group used the criteria of an effective key indicator to reduce the number of potential key indicators from the original brainstormed list.

trustee, was concerned that because he was a volunteer, he did not officially represent the trustee viewpoint. James Oord wanted "a better opportunity to receive direct input from members" as the project progressed; Bill Blair agreed with him.

Stakeholder Team members were also more comfortable when the meeting was being run by the researcher than by the superintendent. Colin Scott commented, for example, that "when the superintendent was in charge of the meeting, I felt a bit irritated...I don't know why...I felt more manipulated, that one meeting". The superintendent himself agreed, stating that:

...to a certain extent it was really helpful to have an outside person. If you are the CEO of a large organization and you're going to evaluate it, it almost looks incestuous unless someone else is there involved in it.

It seemed almost as if it was perfectly acceptable to all that each representative was entitled, as Jason Stone stated, to "represent its interest group's point of view," but if one--say, the superintendent--was put in a position of authority, whereby his influence might be heightened over that of the others, then it appeared to disturb the cooperative dynamics of the operation. The trustee even stated that within the Stakeholder Team "there was a dynamic--and it was more important than the conclusion". The dynamic seemed to evolve towards a sense of group trust, shared involvement, and group decision making.

3. The Four-Goal, Six Attribute Framework

a. Description. A task during the first Stakeholder Team meeting was to examine the Ministry values framework and determine whether it was appropriate as the "ideal" for the Interior District's evaluation. Three sub-groups were formed to independently examine its congruency with the District's existing mission statement. The latter had been developed during the previous year and had been one of the first initiatives of the new management team.

The sub-groups struggled with the task of choosing which of the two frameworks best expressed the values of education in the Interior District. Although the Ministry framework was not particularly contentious in that people violently disagreed with its main thrust, discussions quickly became

centered on semantics, and reading between the lines. Ministry intents were suspect.

During the discussion some differences between the two value statements were noted. The Ministry's was seen to be more categorical, and less holistic than the district's; it seemed to want to compartmentalize learning, as the trustee put it. On the other hand, the District's statement did not clearly distinguish between outcomes (ends, in terms of student learning), and instrumental processes (desired means to contribute to effective learning) like the Ministry statement did. In the end the committee made the tentative decision that "we are...doing what is in the Ministry framework", and that therefore an attempt should be made to use the district's own statement of mission as the values framework guiding the key indicator selection and report writing. This decision was to be revisited as the project progressed.

On two other occasions the decision was debated. In one instance it was noted that the District's mission had little to say about management and accountability, cost-effectiveness, and vocational development. It was also noted that a number of the values in the statement of mission seemed to attract similar indicators; the items did not seem to be quite discreet enough. However, rather than abandon the use of the district statement, Team members made the decision to build indicators of these qualities into to the meaning of the district statements of value. For example, to deal with the concept of vocational development, the district value "to educate students to...experience success" was interpreted as referring, at least in part, to student career development. In this way the intent of the Ministry mandate was maintained, but the ownership the district had built into the development of its own mission statement was preserved.

The topic was again debated during discussions on evaluation report production. To acknowledge the right of the Ministry to guide local education, a decision was made to place at the beginning of the report, a page showing how the District Mission Statement is compatible with the Ministry Four-Goal and Six-Attribute framework (see Appendix E2). Delegation of educational responsibility from the Ministry to the district is thus acknowledged; so too fealty of the district to the Ministry's overall mandate.

b. Critique. Circumstances in the Interior evaluation indicate that the Ministry's framework is viable as a statement of intent. It clarifies the values

that a district must espouse, in their own language and results, in the planning and implementation of educational programs. However, to insist upon its use as the only appropriate wording of those values was, in this case, in the words of the trustee, "a bit negative."

Stakeholder Team members in the Interior District felt obligated to adhere to the Ministry's mandate statement, but wanted the opportunity to express those intents in a way that reflected the priorities and culture of their own local community. Colin Scott, the trustee, commented that "we have to be operating within the Ministry," but sometimes there can "be too much emphasis on that standardization...I think the critical thing that was useful was using our mission statement as opposed to the Ministry framework." Sonja Kocan echoed that comment, stating that: "I thought it was a great idea to use the mission statement. It was something that had been developed in the district for this district." Ownership--i.e., commitment to a concept because one has developed it commensurate with one's own needs and aspirations--was a determining factor in the choice of values framework for the evaluation.

4. Key Indicator Development

a. Description. Much of the Stakeholder Team's time and energy was devoted to the creation, selection, and measurement of key indicators of performance for the Interior School District. Procedures were similar to those employed in the other two districts. Some changes were, however, imposed both by circumstance and design to enhance their development.

A first change was to spend more time educating committee members as to the nature of key indicators and criteria for their development. Stakeholder Team members were given articles to read "as homework"; also, the criteria established for key indicator development as a result of the experience in other districts were discussed at length.

A second change--this a matter of circumstance--was that the two district officers drew up a tentative list of indicators to be presented as the starting point for discussion. This list they developed with the help of the District administrators' group. It was presented to the committee as a series of concepts grouped according to value statements derived from the district mission statement. They (the value statements) were grouped under two headings: "product", referring to all statements that referred to desirable

student development; and "process", referring to the instrumental goals needed to achieve positive student outcomes. During this meeting (which was the one chaired by the superintendent) discussion was open-ended; the group was directed to react to the whole package--and they tended to jump around in concentration and focus. Some group members tended to get frustrated, and solicitation of additional indicators was not too successful.

On the group's next occasion of meeting the process was altered once again. Each framework element was discussed one at a time, and using the brainstorming rules described in the South Coast case study (see above, p. 156; and Appendix E3), additional indicators were induced from the group. Care was taken to enforce the rules and focus the discussion, and not to allow some individuals to dominate. Two conditions were set: first, indicators were broken into two main categories--(1) survey, (qualitative); and (2) non-survey (quantitative and demonstrative). Second, a formal Q-sort was employed to prioritize each and every set of indicators within the two categories of survey and non-survey. Priority judgment was to be based on two things: first, the criteria established for quality key indicators; and second, personal preference. These scores were then tabulated and lists of indicators were circulated to committee members so as to allow them to get reactions from their constituencies (see Appendix E4). "Product" indicators were developed by the group as a whole; "Process" indicators by pairs, reporting back to the main group for feedback.

A third major change in the process was the time spent by the group refining and defining indicators. A full meeting was devoted to this. Quantitative indicators were defined and put in measurable terms. The survey indicators were explained and defined in the presence of the survey contractor so the integrity of intent would be maintained. Demonstration indicators were operationalized as examples, photographs, or samples. In keeping with what had taken place in other districts, there were many different viewpoints as to the appropriate form of data collection, interpretation, and presentation. All in all, four full-day meetings were devoted to key indicator development.

b. Critique. Generally speaking the key indicator development process met with approval from the Stakeholder Team members (see appendix F2). Many of the comments made with regards to dynamics of the committee work describe

the effect of the procedures and processes used. People appreciated input though brainstorming; timelines kept the activity moving forward; and the Q-sort allowed equal input into a final decision. Although sometimes people found the timeline a little rushed, and the Q-sort "constricting", in that people "had" to choose priorities, these procedures generally received a positive rating from the individuals involved.

An improvement in the process was the groundwork done to prepare committee members in terms of understanding. Of ambivalent success was the effort to "pre-develop" indicators--it helped in that examples were provided and groundwork laid, but hindered in that people's creativity was stifled. However, spending time categorizing indicators, refining indicators, and developing the survey in tandem with the key indicator process was seen to be a positive addition to the process. Also, the opportunity to deliberately decide what pictures, or examples would be placed in the evaluation report as "demonstration indicators" met with approval from the group.

One drawback noticed was that the size of the Interior district seemed to make it more difficult to develop indicators that were directly linked to the learning of students. Almost all the measurements were done at a generalizability level that would have to be interpreted at a variety of removes--amongst over thirty schools, and by individuals in district office. In going so the information has to travel through many levels of the educational hierarchy, and has to be interpreted by many actors in the process. This leaves its direct applicability in terms of improving classroom instruction very suspect.

5. Questionnaire/Survey Information

a. Description. In preliminary discussions regarding the Interior District's participation in the IPS project the process of public surveying was discussed extensively. The experiences of questionnaire use in both the Inlet and South Coast district evaluations indicated the need to develop (1) a set of questions that are comparable across constituencies when necessary, but specific otherwise; (2) a set of questions meaningful to the goals of the evaluation, and in the opinion of the Stakeholder Team members important to ask; and (3) an administrative process that was not dependent for its success on the vicissitudes of volunteer expertise and time commitment. Better data

presentation and interpretation capabilities were also desirable. "Ownership" of the process was the desired goal.

Project managers (the researcher, district superintendent, and director) agreed that the best way to ensure these changes was to engage a professional researcher to conduct a poll for the Interior district. This concept was presented to the Stakeholder Team and it agreed, seemingly pleased (particularly the CIDTA president) that control of both the questions and the interpretation of the results would be solely in the hands of the Team

A contract was then let to Windsor Research to conduct a telephone survey of representative samples of the constituents in the Interior School District. Sample sizes were to be determined according to accuracy needs and questionnaire development was to continue apace with key indicator development. Representatives from the company met on two occasions with the committee and submitted the final format of questions to the Team. To ensure some compatibility with the intents of the original Ministry surveys, those documents were used as a guide to assist in wording and intent.

Three surveys were ultimately constructed for the project. One was for parents/non-parents/employers, a telephone survey that identified cluster samples of each group. A second was for secondary school students, to be administered to a sample population drawn from the district as a whole. The third was directed at teachers and non-instructional employees, and was again, a telephone survey (see Appendix E5 for particulars re sample size and administration). The Stakeholder Team had meaningful input into the form and content of the surveys, and on occasion changed the wording and intent of questions, as well as adding/deleting questions according to their priority.

Implementation of the survey was set for January. This created some difficulty as negotiations were in full swing, and emotions were running high. Just prior to its implementation, the superintendent had second thoughts about the potential validity of teacher opinions at this time. He was persuaded to go ahead. Some teachers did, in fact, refuse to answer questions. However, to preserve the integrity of the process, James Oord sent a letter to all teachers asking them to participate in the exercise. He also gave permission for the pollsters to inform teachers of his support. Accordingly, the survey was successfully completed. Results were provided to the committee in late January.

b. Critique. Stakeholder Team members in the Interior School District were quite positive about the success of the survey process. Even the teacher representative commented that the survey was a "positive aspect of the process", in that "one gets a variety of sources of information." Indeed, in the ratings of components of the evaluation process, the accuracy of the Interior district survey was highly rated, completely reversing the trend established in the previous two districts (see Appendix F3).

A key element in the positive response to the questionnaire process was the goal of ownership. For example, Norma Rivette (CUPE) stated that:

...one of the most positive aspects was Frank (contractor)...the knowledge he brought to the committee, and how the surveys were developed to get the answers the committee was looking for for the profile booklet.

There were some criticisms of the survey process, however. The telephone survey was deemed to be too long; Damon Korchuk stated "there needs to be some refinement there." The superintendent would like to have had the results compared to a provincial sample, so they had more meaning. And finally, the results were somewhat tainted by the negotiation issue: as James Oord stated:

I think that the time and political situation took a lot of attention, effort and energy...and I think did skew people's attitudes, specifically teachers...but I think probably also to a certain extent students...

Finally, it should be noted that the survey results were used in the planning process as well as being analysed for inclusion into the public report. An executive summary was presented to the school board and an analysis of potential significance was carried out.

6. Data Collection, Analysis, and Interpretation Processes

a. Description. The Interior evaluation project took place during the school year and this affected data collection--more so than analysis and interpretation--significantly. Data collection assignments were made prior to the early January meeting of the Stakeholder Team. Each person was assigned responsibilities that seemed appropriate to their role. These assignments were then discussed in the committee setting and adjustments were made according

to preference or convenience. Potential data sources were discussed and refinements to the potential measures were made if it was anticipated that only a certain kind of data would be available⁵. Individuals were asked to have their information available for a meeting three weeks later.

Data Analysis and interpretation was done over a two day period in late January. Committee members brought information on quantitative indicators to the session and it was pasted up on the wall, under the appropriate District mission value statement. To ensure significance and meaning, Team members circulated in pairs and evaluated the information according to agreed upon criteria (see Appendix E4). Photos and demonstration indicators were also judged.

As part of the process the survey contractor summarised the main findings of the questionnaires and disseminated the overall results. In order to convert this information into qualitative indicators a small team of two ministry researchers and two Stakeholder Team members analysed the results for significance. They looked for meaningful constituent differences and requested cross-tabular analyses to enhance the information's meaning and understanding.

A second day was spent converting information into prose and visual representations. Once again a small team of Ministry researchers and district staff took on this responsibility, working directly from the raw data created a day earlier.

b. Critique. This component was clearly the least satisfying step of the evaluation project for most Stakeholders. Senja Kocan "thought it was kind of unwieldy--I phone people who had already been phoned a couple of times on other matters". Bill Blair felt that the data gathering could have been better directed by an authority from the central office, who could have accessed the data in a more organized fashion. He even suggested a mini-committee of different individuals charged with this responsibility, and working throughout the duration of the project. And Colin Scott questioned the

⁵ In reviewing the potential data sources it was quickly recognized that a significant amount of information was required with respect to special education programs in the district. It was suggested that it might have been appropriate to have had the Director of Special Services on the Stakeholder Team.

accuracy of the data due to the collection procedures: "When you had stakeholders like myself who are not in the system collecting information, there's going to be a lot of inaccuracy." Damon Korchuk suggested that data collection could have been improved by "making the system more aware of what we are doing, so that when we asked them for information they knew for what purpose the information is going to be used." Some of the principals were reticent to give information, and saw it as a low priority task.

Other comments related to interpretation and significance. Bill Blair once again suggested that it would have been much better to have had the data in advance of the interpretation meeting, "to digest the materials and be able to...have some time for thought." Colin Scott felt that the quality of the data was inevitably weak, because it was culled from a disorganized, disjointed, and never heretofore developed data system. He also recognized that the information was less than satisfactory because it was benchmark data only:

...the whole design was to set something up to build upon...now that we have the categories of information established, we should set our systems for the accurate retention and retrieval of information.

Jason Stone, the superintendent, agreed, although in a somewhat qualified sense. He stated that there should be better data collection procedures "if we're going to use it"; if the evaluation is not a "one-shot" event.

7. The Evaluation Reports

a. Description. Before the project began it was agreed that the Interior District evaluation report would take two forms: (1) a video outlining the current state of education, in the form of key indicators--containing footage that brought the district vision and its related statistics alive; and (2) an evaluation booklet, modeled after the Inlet and South Coast versions, that would be used to summarize the state of all key indicators chosen to characterize the district's mission statement.

The Interior District administration team had been motivated to take part in the District evaluation because they saw potential for enhancing the effectiveness of their "video" annual report. The previous year's attempt had been criticized for being too positive. The superintendent was taken by the idea that a video reporting both the strengths and weaknesses of the district as

revealed from an evaluation, could be used to engender public support for improvement of the system. The researcher was intrigued by the concept because it gave the opportunity of communicating evaluation results in a format that would appeal to a second learning style.

A second virtue of the video was that it allowed a fundamental component of effective education--the quality of human interaction--to be communicated at the same time one was providing quantitative data attesting to its accomplishments (or lack of them). In this way the vision of a system constructed to serve the needs of students could be properly communicated.

Even though the formation of a video and the creation of a printed evaluation booklet was decided at the outset, details as to intended audience, length, format, and purpose had to be worked out through discussion at the committee level. Sid Pollack's responsibility was to convert the evaluation information into a video: he was given artistic licence here, receiving very little guidance from either the Team or the researcher. However, the committee had much to say about the format of the printed report. There was significant discussion about whether there should be "two separate booklets...one for employees, one for the community"⁶. The researcher suggested that the purpose of the evaluation exercise should determine the nature of the report; i.e. if one wishes to use the evaluation information for internal planning, then a report should be written accordingly--in the language and style that appealed to educators. The same argument applied to a report for the purpose of engendering public support for education.

The committee finally made the decision to produce a report in the form of a public document. It was to complement the video--in other words, the video would be used to introduce the concept of evaluation and give a sample of its intended use, and the report was to serve as a vehicle for detailed discussion of the state of education in the district. Ministry researchers spent two days in the district establishing the general format and content of the report, and then returned to home base to steer the production of the booklet towards the intended deadline.

⁶ Damon Korchuk, Director of Instruction--Communications and Administration.

Report production for both the video and the booklet took five weeks. Committee approval was forthcoming on the general readability level of the document, and each page--prose and graphs--were faxed to district for editing, comments for improvement, and final approval. Professional attention was given to graphics, layout, and the use of color to enhance information presentation. Pictures were included as demonstration indicators. Research quotations were used to characterize desired goals or outcomes. Most of the work was completed by the Ministry team, but the district gave final approval to each and every entry in the booklet (even though at times this created time problems for production).

Video production was more problematic than the printed publication. A first version was produced that was not successful in capturing the message envisaged by either the superintendent or the researcher. At this point a professional script writer and producer was hired. The final product expressed the role evaluation would play in district improvement more effectively and was, therefore, commensurate with the intent of the project. Video cost was approximately \$15,000 and printed report production about \$8000. Half the costs of each were assumed by the Ministry of Education, half by the district.

b. Critique. Stakeholder Team members provided a variety of perspectives on the success of the procedures and processes used to develop the evaluation report. Strengths relate to the efficiency of the process, the objectivity of the report(s) and the ability of the report(s) to represent the uniqueness of the district. Weaknesses are in the areas of resource utilization, information presentation, size limitations, audience specificity, and expertise requirements.

The superintendent spoke highly of the willingness of Ministry personnel to work in a "partnership effort together" during the production of the report. The Ministry took on the role of providing expertise in information presentation--i.e., graphing, prose and layout--while district personnel proof read and examined the presentation to ensure the integrity of the information. This assignment of responsibility ensured district control of the result, but provided expertise that would otherwise have been inaccessible to them. The written report was completed quickly and efficiently on schedule. The result was a report that Jason Stone called a quality product:

...I'm proud to have that out there in offices...regardless that its got some negative stuff in it...its done with class, and it looks like it represents a \$70 million operation.

It also preserved the uniqueness of the school district. All Stakeholder Team members agreed, and Bill Biair expressed it vigourously when he stated that "the process allowed for the uniqueness of the district to be posted up front...in the final report, some sections...really delineated that." Even though video production was more problematic, the final result expressed the characteristics unique to the district.

Accuracy of reporting was a second strength of the two reports. Stakeholder Team members were pleased that the reports contained both positive and negative information, and tended to give a "fair" treatment of it: not only both positive and negative information, but also a report that "evaluates all facets and aspects of the school system, from the point of view of the groups that are immediately or indirectly affected."⁷ In so doing the report gave "a clear view of future directions," and Bill Blair stated, it "is a practical tool both at the school level, and for school district management." In spite of these strengths there were a number of areas for potential improvement. Bill Blair thought a number of "the graphics were far too complex in design...we need to be as simplistic as possible." The pictures were a positive component of the reports, but "they could be more pointed as they relate to each of the key indicators." He suggested that more time should have been provided for their taking.

Some teachers felt that despite the reports' quality, they "left out a number of important components relating to the classroom". They also established "some predetermined ideas." The former was inevitable in a finite document, and was also alluded to as a problem in forming key indicators in a large district. The latter was deliberate, as the point of the exercise was for the district to create a series of key indicators of performance that are "predetermined" in that they are agreed upon measures that will be used to monitor the quality of educational performance in the near future. This

⁷ James Oord, President of CIDTA.

comment actually reveals misunderstanding on behalf of some teachers as to the purposes of the project.

A third potential weakness was the attempt to appeal to more than one audience through the production of reports primarily targeted at the public audience. As Damon Korchuk stated, "at one time we were looking at doing two separate booklets...one for employees, one for the community. The next thing you knew we tried to do both in one." Although the decision was bolstered by the production of a video, there is some doubt as to whether the level of generalizability of information required for public consumption is meaningful at the professional level. The Stakeholder Team itself had split opinions on this, characterizing the report actually as more effective for management purposes than public information purposes.

A final criticism of the report production process relates to resource use and accessible expertise. The Interior District had very limited funds to facilitate the production of the reports. In addition, they could not have done it without "some technical expertise" from the Ministry. Jason Stone commented, for example, that "we really appreciated having somebody who could do graphics and that kind of...(work)." The District could not have accomplished the project on their own; in this sense the process and procedures are not immediately transferable to the District setting.

Applications of the Interior District Evaluation

Description

There were a number of different strategies developed within the Interior School District to use the evaluation's information. These strategies revealed a two-prong thrust: one directed at accountability and gain of professional and public support for education (i.e., public relations); and the second at the use of the evaluation information for planning purposes (i.e., policy making).

The Interior District took advantage of the opportunity to examine strategies for the use of evaluation information employed in other districts. They did the following activities in the area of accountability and gain of public and professional support for education:

1. During education week, the video was run continuously in a number of shopping malls, and a manned booth was located

nearby; individuals were encouraged to take an evaluation report for personal reading.

2. Evaluation Booklets were placed in the waiting rooms of dentists and doctors throughout the school district.
3. Members of the management team took the video and booklets to service club meetings, using the video as a lead in to analysis of the book, followed by discussion of the District's educational performance;
4. The superintendent presented the video and booklet at a Chamber of Commerce meeting in Appleton.
5. Copies of the booklet were sent to each school in the district and the superintendent "talked to principals about what they should be doing with parent groups and the staff...Every school will be using it".
6. Parents groups in each school used the video and booklet as the topic of a parent meeting.
7. Media contacts were made, including a television program using clips from the video and interviews with the Director of Communications.
8. The video has been placed in the foyer of the District office and is played for visitors while they are waiting.
9. Evaluation reports have been disseminated to real estate companies so they can be distributed to individuals moving into the Interior area.

Use of the evaluation report for these purposes was not restricted to district office staff. Principals and the parent representative on the Stakeholder Team were involved in chairing meetings to discuss the project.

The evaluation information--both that contained in the formal reports, and the more detailed information gathered through the survey--was used to improve planning in the district. This happened both formally and informally. One formal application was to supply the District planning committee that had already been established as part of the District's goal setting process with evaluation reports. The evaluation was seen as providing additional information to assist with that goal.

A second formal application was use of the evaluation information in budget planning. Colin Scott commented, for example, that:

...the increases that we approved above the sort of maintenance areas were really related...to the evaluation...the whole (budget) section on communications was a result of this evaluation.

Informal evaluation information applications derived from reliance on the strategy to allow each individual to determine the significance of the information, and act accordingly. Louise Loewen, for example, commented that:

...teachers have been reading it...I think that what is going to happen with the teachers is that they are going to look at fixing physics, but at the same time they are going to go and fix things up that weren't caught this time so the next time they are not around to get caught.

Although some teachers criticized the report as a bit of a "whitewash", the very fact that information was collected and reported and might be again could well spur improvement.

A second informal application of the evaluation information was with respect to Stakeholder Team understanding and commitment to district improvement. Many commented that the process identified "obvious" improvement steps; as Bill Blair stated, "there are some perceptions from the student or public view that really indicate that we have some work to do, and it's our responsibility now to take that information and begin work on it." Each one stated they were more committed to district improvement after rather than before the project. The student representative stated how "she learned a lot", mostly in terms of broadening her understanding both of district operations and the needs of her fellow students district-wide. Jason Stone commented on how other members of the stakeholder team were led to a clearer picture of the district, and to a greater appreciation of both the complexity and size of its operation.

Critique

Stakeholder Team members provided a great deal of insight into the effectiveness of the evaluation follow-up procedures in the Interior School district. To a great extent they considered the evaluation to have had a significant effect on district direction. There was enough evidence to show, for example, that its formal use enhanced the planning process in the district.

Evidence also exists to demonstrate that its informal use by Stakeholder Team members, teachers, and parent groups (1) improved their sense of district direction; (2) identified areas of improvement they themselves could affect; and (3) motivated them to commit themselves to district improvement. However, a number of suggestions as to how to improve follow-up procedures were also suggested.

A strength of the evaluation application process was its affect on the budget and planning processes. The trustee stated that it "was very useful in our budget process." He also stated that "it affects decisions that I'm making...in terms of trying to keep the issues in front of the board...": "the whole evaluation has provided a frame of reference for me." The evaluation provided a means to enable trustees to identify important issues of policy-making⁸; prior to its use "they...(the trustees)...were quite detached from that, not by design but by default."

A second strength of the application procedures--probably due as much to the structure of the whole evaluation process as much as follow-up itself--was its ability to enhance understanding and acceptance of the District's mission. This was seen by many Stakeholder Team members as fundamental to effective planning. Bill Blair, for example, was of the opinion that the evaluation process "gives us some baseline information from which to develop and plan for the long term." Arguing that knowledge of the district's strengths and weaknesses will be the goad to improvement, he stated that the report "definitely" gave a "clear view as to what they are" in terms of "a total district picture." He was of the opinion that the evaluation gave substance to the efforts the district had made to develop a vision, or mission, by providing "data that will allow you to grow and develop an appreciation of the professional statements that have been made about the educational system." Damon Korchuk agreed, stating that:

...what this will do is give us some data as to how well we are achieving the mission statement, and maybe what it will do is rekindle and bring forward the whole concept of having a mission... its been excellent for us that way.

⁸ Colin Scott, the trustee, commented for example how the evaluation project had "made him aware" of the district's drop-out rates. This became an important issue to be dealt with in planning.

A third strength also relates to the effect the follow-up strategy had on the planning process. It was alluded to in the discussion of budget. The evaluation was helpful in establishing what the superintendent would call "growth goals". A potential role in helping express "maintenance goals"--i.e., goals that reflect the base-line expenditures of the district--in terms of measurable outcomes, was seen as a desirable, but as yet unable to be realized. This limitation will be overcome if the process becomes systematic, and is built into the administrative fabric of the district. This points out that the IPS is a very fledgling process, and its true promise--in the eyes of the Stakeholder Team at least--can only be realized through ongoing, systematic use.

James Oord suggested that the planning process could be enhanced by building a cooperative response strategy to the evaluation reports. The current strategy involved issuing each teacher in the district with a copy of the report. Therefore, "each person look(ed) at the report from his or her own perspective and look(ed) for certain things; they want(ed) to see work that they have been involved in reflected in the report." The teacher is, in other words, looking for ownership. Without it, personal interest wanes. A cooperative response strategy is a mechanism to create ownership, at the "staff level, teaching level,...(and in)...local specialist associations groups,...and parent groups."

Finally, one particularly insightful comment about the evaluation's application--as a process rather than just as a source of information--came from Colin Scott, the trustee. He described, in a sense, a control versus a treatment experimental situation. He compared this evaluation with the previous one conducted in 1986, after which time the management of the district went through turmoil:

Our district had gone through, with the other superintendent, a very stressful time...we ended up doing an internal evaluation, and I had extreme difficulty with that because it became very personality oriented.

Big systems like this, no matter who is in the key position, will always have stresses and strains. This process allows ventilation and expression of meaningful information. If people had wanted to they could have expressed any concern, but they would have been forced to express it in a constructive way, in terms of results...For example, if caring staff is being impeded because of policies, or structures, it

would have been possible to have aired that, in a way that would have brought informal pressure...so the reason I think it should be used systematically is that it supports the informal communications system--gives it legitimacy--by focusing it on constructive results.

This comment implies that informal applications of evaluation information are possibly as important as the formal. In some ways Colin Scott suggests they may redress potential ethical problems that cannot be sorted out through a formal process, or that might require the construction of a very cumbersome formal process. This insight will be considered further in the final chapter of the dissertation.

In conclusion it must be noted that efforts to apply evaluation information in the Interior School District were quite intensive. It should also be noted that the process allowed for the dissemination of both formal and informal information, and that both forms seemed to have a role to play in reaching the evaluation's objectives. Finally, it should be commented that the evaluation process seemed to create ownership of the results, and this appeared to be a factor enhancing its use.

Summary

The Interior School District Evaluation can, in light of the evidence, be cautiously termed a success. There is ample evidence to show that it did have some success with respect to its anticipated goals of improved accountability, gain of public support for education, and enhancing long range planning. Yet most of its potential seems to be still in the future. Without ongoing processes--including a better groundwork process, improved school based data systems, more effective follow-up procedures (particularly with regards to receiving reactions to the evaluation), and solutions to the resource problem and expertise problems--its future potential is almost nil. The true test of its value--whether or not it is actually built into the administrative structures of the district as a systematic process remains to be seen.

CHAPTER VII

Meta-Evaluation of the Information Profile System

In keeping with the design presented in Chapter III, the IPS will be meta-evaluated (i.e., across case studies) in two main areas: methodology; and effectiveness. The design outlined in Chapter III determined the data and analytical processes employed to conduct the meta-evaluation.

Meta-Evaluation of the IPS Methodology

Four questions, derived from the IPS's initial pilot study, were created to guide this stage of the meta-evaluation. They are:

1. How acceptable is the Four-goal Six-attribute framework in representing the overall purposes and methods of the education system?
2. What indicators, other than those currently in use, are appropriate measures for each goal and attribute?
3. What changes should be made to the questionnaires and reporting-out formats to make the IPS more effective? and
4. What processes and procedures were valuable in focussing evaluation team members on the task of evaluation?

1. The Four-Goal, Six-Attribute Framework

In each district evaluation stakeholders were encouraged to use the Ministry values framework. During the first pilot the four goals and six attributes were used as descriptors of the qualities of an ideal education system. Reactions were almost blase. Educators seemed to take the framework for granted; few suggestions were forthcoming for improvement. Yet some dissatisfaction was expressed, particularly with regard to the attributes. A purpose of this study was to critically examine its acceptance and applicability from two viewpoints: (1) What role did a values structure play in the evaluation; and (2) Were the values embedded in this particular framework representative of the three districts participating in the pilot?

Patterns revealed in the analysis of the three case studies show that the 4-Goal 6-Attribute framework appeared to play two practical roles during the evaluation process. First, it was accepted by all three districts as a statement of "ideal" values deemed fundamental to a public education system. Therefore, it provided a frame of reference, or set of philosophical criteria against which each district could determine its performance level. Two of the three districts (Inlet and South Coast) used the 4-Goal and 6-Attribute framework with minimal modification, and the third (Interior) was quite content to defer to it as a "statement of provincial intent" to which the district was obligated. In the last case the Stakeholder Team utilized their own mission statement in conducting the evaluation, but not because of disagreement with the content of the Ministry framework: they felt that the district's own mission statement expressed the same values. But, because it had been developed locally, its use ensured district autonomy and ownership of the evaluation process.

A second role the framework played in the evaluation process was to focus evaluators on student outcomes. Cooley and Lohnes (1976) emphasized the importance of this quality in formative evaluations employed by educational administrators. The four goals encouraged the development of key indicators that measured system effectiveness in terms of student performance. The student representative on the Interior Stakeholder Team expressed it this way:

The most important thing in this whole process was to find out how our students were doing. Not how schools are doing compared to one another, but how students are doing, and learning, and functioning...that I definitely felt was the whole overall goal and I felt we were successful in that.

Stakeholders in the Inlet evaluation agreed, commenting how the framework had dignified the efforts teachers were making to assist student social and personal development, and how it personally focused them on meeting these needs. In the case of the Interior district, the role was complicated a little when the District chose to use its own mission statement. However, a focus on student outcomes was achieved by extracting statements from the mission that were similar in intent to the Ministry's four goals (see Appendix F11 for stakeholder opinions with regard to the focus on student outcomes). There is also evidence to show that the information on student performance so

gathered also became the focus of planning in at least two of the three districts¹.

But how about the wording and substance of the 4-Goal 6-Attribute Framework? Analysis of District reactions indicate that stakeholders found it reasonably accurate (See Appendix F1). Almost all the concepts inherent in framework definitions were acceptable in each district. What did become a source of debate were whether titles were appropriate, and whether separation of some of these intents into discreet packages was appropriately done, or even academically defensible. People in both the Inlet and South Coast Districts found the distinction between social and human development somewhat artificial. Stakeholders in the Interior district found the attempt to compartmentalize these goals a little restricting--and contrary to the holistic nature of student development.

Stakeholders in the Interior District were the only individuals to truly debate both the wording and substance of the values framework. Wording only became an issue because ownership of the District's mission statement was important to Stakeholder Team members. When discrepancies were noted in substance, i.e., when analysis of the District mission statement revealed that it had no overt statement relating to a goal or attribute, Stakeholder Team members compensated by developing indicators of student performance that measured this value, and placed them under a related mission statement phrase. This was the case with regard to the issue of vocational development: Stakeholder Team members developed indicators of student performance and placed them under the rubric of "developing students to meet the challenge of change." Similar measures were taken to ensure aspects of cost-effectiveness and management and accountability.

The three case studies suggest a number of possible changes to improve the wording and substance of the 4-goal and 6-attribute framework. A first is to reexamine the validity of the human and social development distinction. Confusion arises over the categories (i.e., whether it should be human

¹ In the Inlet District, efforts were made to address English as a Second Language (ESL) needs, and to find out why students did not perceive the schools to be as caring as District officials hoped; in the Interior District, drop-out rates became an issue for the trustee representative on the Stakeholder Team. The level of teacher caring revealed by the student survey was also identified as an area of improvement in the district video.

development, or individual development) and difficulty in distinguishing between measures of each needs to be resolved². Secondly, the appropriateness of the term "cost-effectiveness" and its definition needs to be reconsidered. Two of the three districts agreed that the value of "fiscal responsibility" is a more realistic expectation for the education system; the other adopted key indicators reflecting fiscal responsibility rather than cost-effectiveness. Third, if the framework is intended to be a values statement expressing ideal "product" and "process" values for the education system, then the terms "management and accountability" and "public satisfaction" need to be renamed and redefined consistent with this intent. The terms "rationality"--to indicate a quality desired in all organizational decision making, and certainly a meta-value in education--"quality management", and accountability are suggested.

In conclusion it is recommended that the 4-goal and 6-attribute framework (in a revised format) remain as the recommended values statement for the IPS. It focuses on student performance, it represents ideal values against which key indicators can be measured, and it expresses these values in a language that will encourage fruitful educational discussion. In terms of wording and substance the 4-Goal and 6-Attribute Framework is general enough to adapt to individual situations (e.g. by allowing for each district to choose its own key indicators to express the goal or attribute). In cases where the district has developed its own mission statement, it provides a framework against which it can be critically examined and compared. Within the loosely-coupled nature of the B.C. education system the values framework can act as a unifying force between school districts without confining creativity or local expression.

2. Key Indicator Measurements of the Goals and Attributes

Experience with the first Ministry pilot indicated that districts used information from the Ministry's central data file and from the surveys as key indicators of performance. The Ministry's data files contained no quantitative data on a number of the goals and attributes, and survey questions were not

² The Government of British Columbia, in responding to the Royal Commission on Education, resolved these issues through the creation of a new Mandate Statement for Education in February, 1989. In this statement the goals of social and human development are combined (see Appendix F14).

constructed so as to be necessarily consistent with the goal and attribute framework. Two purposes of each case study were, therefore, to see if: (1) quantitative measures of each goal or attribute could be developed; and (2) there were qualitative measures that could be expressed as specific survey questions for each goal or attribute.

To answer these questions the IPS was redesigned so as to give stakeholders within districts the opportunity to design such measures. The intent was twofold: first, to see if efforts to develop measures of educational performance in the area of social and human development of students would, in fact, be acceptable to stakeholders; and second, to see if the measures so chosen--in these goal areas and others--would possibly be amenable to standardization.

Somewhat surprisingly there was little reluctance on behalf of stakeholders in any of the three districts to develop measures of educational performance. Indeed, stakeholders responded in a manner that could be termed "moderately enthusiastic." Patterns of response in the three districts showed that in the Inlet and Interior districts Stakeholder Team members were more positive about key indicator use than stakeholders in the South Coast district; and whereas the concept of key indicator use received a reasonable endorsement, the consensual perception of the accuracy of the resulting key indicators was supported to a lesser degree (see Appendix F2). This is likely due to the fact that the quality of data, the process of data interpretation, and compromises made to report it made it difficult to realize the potential of a number of the indicators.

It is also interesting to note that this same pattern was repeated with regard to examination data and other indicator information provided by the Ministry. The district that was least supportive of this information was the Inlet district; both South Coast and Interior districts reacted in a very positive way (see Appendix F2). Possible reasons for this state of affairs are threefold. First, the Stakeholder Team in the Inlet district had three teachers on it; in the other two districts only one teacher representative was present. Second, the data for these key indicators were credible, whereas the data for the district-developed indicators were somewhat suspect. Third, districts were allowed to interpret and represent this information as they saw fit, and therefore ownership of the results was created. This was the case in all districts.

A third pattern noted was that teachers, as a group, tended to be less enthusiastic about key indicator measures than did other stakeholders. Surprisingly enough the district indicators were no more popular with educators than were the Ministry indicators. This condition seems to belie the enthusiasm teachers evinced at the beginning of each of the projects with respect to expanding measures of performance into the social and personal development areas. However, this is possibly due to the quality of the data available for these indicators rather than the concept itself; district indicators were not anywhere near as statistically sound as were the Ministry ones. Another possible explanation is that it may reflect discomfort with the whole exercise, as a reaction to BCTF rumblings re accountability efforts.

Another pattern relates to the effect of the processes used to develop key indicators. Brainstorming, reliance on intuitive judgement, the use of the Q-sort for priority setting, etc.--helped create a sense of ownership and originality for each district's set of indicators. The result is a list of indicators that is important to each individual district. And although a number of these indicators--eg. suspension rates--were used by all three districts, the process that resulted in ownership also resulted in each being characterized in a manner unique to that district.

Experience with brainstorming, the associated debate, and then the processes of data collection, analysis and interpretation, demonstrates that standardization may not in fact be either desirable or possible. Each district emphasized the necessity for its key indicators to express its unique context. Each district wished to surround the statistical data with pictures and information that expressed that context. Each district gathered the statistical information then expressed it in prose and visual ways that gave it different meaning³. Each district tended to react to the other districts' evaluation reports and key indicator representations as "incorrect", or "not how we would do it"; each district strove to improve the quality of information so to make its meaning clearer, fairer, and self-evident. And finally, each district wanted control of the questions in the questionnaire, wanted to choose the sub-

³ Even "given" indicators, such as provincial examination scores and reading assessment tests (as measures of intellectual development) were expressed differently from district to district. Efforts to improve the quality of expression demonstrated dissatisfaction with the "fairness" of the message contained within the evaluation booklet.

samples that were important to them, and wanted to analyse the data according to their own needs. All of these wishes suggest standardization is not desirable; the problems of information management suggest that it may not be possible.

However, at times a countervailing wish was also present. Stakeholder Team members used comparison to create perspective. The Interior superintendent, for example, wanted to compare his district's survey results to a provincial sample; and South Coast stakeholders were extremely frustrated at the inability to compare results on the questionnaire from one constituent group to another. These observations suggest that if the goal of standardization and the goal of ownership are both to be achieved, a provincial process of indicator development should be attempted. If a provincial stakeholder group, representing all major education interests was formed, then a set of standardized indicators, "owned" by all constituents, might be acceptable to districts. Each district could then be free to augment this standard set of indicators with those of their own choosing.

Patterns revealed through analysis of the key indicator development process in each district suggest the following criteria to describe effective key indicators:

Key indicators should:

1. be representations of education performance, either quantitative (i.e., hard data, such as test scores, student participation rates); qualitative (perception data through surveys); or through demonstration (pictures, sample work, sample expectations).
2. be significant, meaningful outcome representations of a goal or attribute to the respective stakeholders.
3. focus on student performance wherever possible.
4. reflect changes in educational action through fluctuation in score; or by demonstrating agreed upon standards of quality represented by a score, or accepted practice.
5. represent, either singly or in concert, the performance of the whole range of students.
6. be such that changes in score can be clearly attributed to changes in educational action.

The experiences of this research suggest also that key indicator selections for each goal and attribute are most powerful when a quantitative, a

qualitative, and a demonstration indicator are used to express them. Examples can be found in the evaluation reports of all the districts studied. Key indicators are relatively meaningless as single year statistics; they are much more meaningful expressed as trends over time.

Two other factors related to key indicators should be mentioned. Districts struggled with interpretation of key indicator definition and information. However, the opportunity does exist to define a meaningful data base upon which measurements of educational performance can be made. It is also suggested that a further stage of development would be to determine expectations of performance for each indicator⁴ so that performance judgments could be based on clearly defined criteria.

Finally, it should be noted that the same situational difference that gives rise to "unique" indicators for each district could be as much a function of time as context. The goal of the IPS model is to evaluate the district's effectiveness in terms of realizing its ideal values, not the specific indicators themselves. Therefore the inevitable trend to institutionalization of indicators must be resisted. Indicators are time and situation specific. Their lifespan as meaningful extrapolations of the framework values is limited accordingly. Intellectual development measures that are appropriate today may not be tomorrow; at the very least the expectations of performance may alter. It is vital, therefore, that the key indicator development process be made systematic, and that the process remain inductive and goal-free (Scriven, 1974). In this way the indicator pool can grow and fluctuate. Indicators that are consistently chosen through such a process can become--as long as they are durable--measures of system performance. In this instance the measure will not drive action; rather, educators will be motivated to improve the original educational value, as no one will be able to predict, except for the very short term, what indicators will be chosen to measure the value's presence.

⁴ Note the term "expectations" is used here, and not "standards". Conley (1988) points out that standards create practical difficulties because they require--demand--consequential action if they are not met. In a "public" education system, these consequences may not be politically possible or desirable; therefore the word expectations is both safer and more appropriate to the realities of education politics.

3. Questionnaire/Survey Procedures and Evaluation Reports

Survey procedures and the methods used to report evaluation information varied significantly district to district. Patterns reveal potential areas for improvement.

a. Questionnaires. The biggest variation in stakeholder reaction to the IPS's procedures and processes occurred with regard to surveys of constituents. In the Inlet and South Coast evaluations, the survey procedures and processes occasioned the greatest criticism and the lowest ratings in terms of accuracy; in the Interior evaluation, the trend was reversed (see Appendix F3). Why such a difference?

The developers of the Interior evaluation had the opportunity of learning from the other two projects. Those experiences suggested two areas for consideration. First, surveys of public opinion of education were complex, time-consuming, and demanding of specific knowledge and skills. Second, survey questions that do not reflect the needs, aspirations, or interests of district stakeholders are not treated with the same interest or seriousness as those that do. "Imposed" questionnaires from the Ministry are suspect as to intent and meaning. Thus managers of the Interior evaluation had two choices: they could (1) manage the survey process themselves, which would require significant amounts of time both for training and administration; or, (2) hire a contractor to manage this component of the evaluation.

In choosing the latter option, Interior District evaluators were able to correct many of the weaknesses pointed out by the stakeholders in the Inlet and South Coast evaluations. Analysis of the three experiences leads to the following suggestions:

1. Survey questions should be developed as part of the overall indicator development process.
2. Sampling of the overall population should be done, rather than surveying indiscriminately.
3. Cluster, or stratified sampling should be a function of the dynamics within the district--i.e., dependent on which subgroup's opinions are vital to understanding.
4. Survey questions should reflect, as much as possible, the criteria of a good indicator.

5. Efforts should be taken to ensure that each sub-sample is asked a parallel version of a question that is best understood through comparison of response patterns.
6. The method of administration (i.e., telephone survey, self-administered, etc.; sample size) should be determined by predictions of success and desired accuracy of results.

The scope of the survey will of course be determined by the resources available to the district. Issues such as survey length, wording, etc. are best worked out in partnership between the contractor and the district stakeholder group. If the link between the survey questions and the values framework is determined in advance, analysis and interpretation activities are facilitated. In terms of timing, it is clear that administering surveys at year-end or year-beginning is not a good idea; neither is administering the survey during negotiation time. The survey is probably best administered during the months of April and May. A final suggestion is that the survey ask respondents to rate the importance of a question and its inherent value, as well as rating the district's performance on that value. This allows the district to do a cross-break analysis which quickly distinguishes between high and low priority issues (Alberty & Mihalik, 1989).

b. Reports. Analysis of the patterns revealed with respect to report writing and information use suggests a number of strengths and weaknesses with this component of the IPS. Experiences suggest that one report aimed at all audiences may be efficient but its impact is limited. In the Inlet and the Interior evaluation the decision was made to pitch a single report at the level of both the public and the teaching audiences. The result was that the teaching audience perceived the document to contain elements of a "whitewash" (i.e., a public relations exercise); and the public found it too detailed and lengthy to read without some guidance or structure to encourage understanding and response. The South Coast newspaper insert received much wider circulation and was easier to digest. In the Interior case study the combination of video and booklet enhanced the quality of communication to teachers, businessmen, and parent groups. Further investigation about the utility of video as a meaningful evaluation reporting mechanism should take place.

Suggested technical changes in the reporting of evaluation results should continue with respect to graphics, layout, and prose variations (i.e., readability level and comprehensiveness). Stakeholder reaction to the reports indicated that progress was made as each report was developed, but that many visuals still ended up being difficult to understand or interpret. Clearer, more insightful visual representations of data (including the use of color) are to be desired. The impact these technical changes have on the integrity and meaning of data must also be investigated further. The strong possibility exists that in the process of generalizing information, it becomes impossible to convey the integrity of the specific meaning of the data. It is also possible, as pointed out in the South Coast case study, that there is no inherent truth, or intrinsic integrity to data--it may simply be a reflection of perspective, individual or shared. This assertion is particularly important in light of the comment made earlier about how important the management of meaning is to the administrator. There is evidence in this study that shows that administrative expertise in these areas is lacking, both in a technical as well as an ethical sense. Indeed, part of the problem may be in that administrators see information management as a technical matter, and not an ethical one.

Finally, evidence in the three case studies indicates that when the purpose of the report is predetermined and specified, there is a clearer likelihood of that use being realized. In each district the project's prime goal (either explicitly stated or implicit in action) was to assist the district with the maintenance and gain of public support for education. A planning role was also acknowledged, but it was not understood or articulated strongly enough to be distinguished from the other. As a result the reports reflected the first purpose more than the second. And, without a strategy developed in advance for the planning use, the public relations role for the report took precedence.

Further studies should examine the impact of evaluation reports specifically aimed at the planning purpose. In none of the three districts was an attempt made to produce an evaluation report that was more than an a "balance sheet" report--i.e., a "picture" of the system in a point of time. An alternative reporting process, of a profit and loss nature, whereby monthly updates of certain indicators are disseminated and collected in a "master planning booklet" might allow for more immediate and meaningful application in the planning process.

4. Focusing Stakeholders on the Task of Evaluation

Evaluation is a process whereby individuals or groups make judgments as to the value of an object--or opinion--as measured against an ideal (the three stages of holistic evaluation are described on page 85). Evaluation can either be internal (i.e., intuitive, according to one's own ideals and perspectives) or external (i.e., consensual or arbitrated, comparing agreed upon ideals with cooperatively or combatively resolved multiperspectives). The purpose of the IPS evaluation model was to develop an explicit, *consensual* evaluation model for school districts that would be based on the intuitive evaluation skills of stakeholders. In so doing it was hoped that the use of evaluation information would be enhanced, and administrative goals achieved.

As such, it can be argued that all the processes and procedures employed in the model focused stakeholders on the task of evaluation. All major tasks were constructed so as to employ the comparison of evidence against some ideal criteria, and then to use individual intuitive judgments (multiperspectives) to create negotiated group judgments. In key indicator development, for example, personal evidence was the source of the initial measures of performance. Evaluation skills were again employed in measuring each suggested indicator against the criteria for an effective indicator; and again in resolution of differences of opinion through the use of the Q-sort. In interpreting data, intuitive judgments were used further in conjunction with color-coding criteria (see Appendix C4 for an example). Even report production was nothing more than a series of orchestrated steps in judgment; stakeholder members chose appropriate visual and prose representations of information, and relied on either consensual or intuitive judgement at the various stages of editing and proofing.

Devices such as brainstorming rules, key indicator criteria, the Q-sort, data analysis criteria--used in concert with the Stakeholder Collaborative concept--were deliberate attempts to foster team evaluation skills. The public surveys required that respondents used evaluation expertise, and interpretation of the results required that stakeholders used the same skills.

Interpretation of the written reports relied on the reader's evaluation skills. They were designed to provide people with: (1) a definition of the values framework; (2) information re key indicator performance, including

pictures and examples; and (3) quotations from research to link the ideals with key indicator choice. In so doing they encouraged the reader to make his own judgments as to how well a particular value is characterized in the school district (this was noted in the comments of Dan Roebottom in South Coast, and James Oord in the Interior district).

Stakeholder members in all three districts generally found the procedures and processes in the evaluation functional and efficient (see Appendix F4). They also felt the process was not overly rule-bound (less so in the South Coast and Inlet Districts, likely because of the imposition of Ministry questionnaires on the process). However, the one component of the evaluation project that all districts found cumbersome and inefficient was the data gathering process. This was due to the paucity of data available at either the school or district level, and the fact that no system was in place in any of the districts to collect important information. This was exacerbated in the Inlet and South Coast districts because of the project's timing.

Further activities that would enhance the use of evaluation skills would be to develop follow-up strategies that encourage individuals to develop and share their suggestions for district improvement with one another. If the whole process was then employed systematically, i.e., every two or three years in a cyclical fashion, the District administration could be constantly informed through a structured, consensual evaluation process. Determinations as to the effectiveness of such a process are beyond the scope of this study and are subject to further investigation.

The Stakeholder Collaborative Approach

A major change to the IPS methodology was not properly dealt with in answering the four questions posed on p. 200. It has to do with the use of the stakeholder collaborative team to conduct the evaluations. This change should be maintained. The case studies have suggested a number of guidelines to follow when building such a team.

First there should be clear identification of the main interest groups that will facilitate overall achievement of the evaluation's goals, or that will be affected by its outcomes. Each group could then be given the opportunity to determine their own representation. In future evaluations it is suggested that the superintendent, an administrator, a student, geographical representatives

from major centres, teachers from both elementary and high school, parent(s), and representatives from the support union be placed on the stakeholder committee. In addition, age and gender should be a factor in selection, and consideration should be given to adding a non-parent and/or employer to the committee. However, the question must be asked, "How does one decide which groups to include, how many, and who chooses"? There is, of course, no external reference point for answering these questions. The research indicates that a committee of eight to 15 people is a manageable group to work with, and composition of that committee ultimately should reflect the major political interest groups that the administrator must deal with. The creation of a stakeholder group provides input from various interest groups, but also demands that administrators possess very high levels of group management skills.

Prior to selection each group should insist that proper groundwork be done so the purposes of the evaluation are clearly defined. Individuals chosen should be knowledgeable, have good interpersonal skills, be creative and flexible, and be possessed of a positive nature. They should come as a representative, not as a delegate; in other words, they must be free to make decisions in a group setting without being hamstrung by the need to report back on all contentious issues before being allowed to make decisions. Mechanisms should be constructed to allow stakeholders to report back to their parent groups while the evaluation is in progress. In addition, a role explaining the evaluation results and gathering reactions to the evaluation report might also be considered.

Finally, it is suggested that the group's activities be facilitated by an outside evaluator. Credibility is provided for the evaluation, and role-equality is maintained within the stakeholder group.

In conclusion, these concepts and ideas have led to the model and timeline for the evaluation process outlined in Appendix F13.

Meta-Evaluation of the IPS's Effectiveness

The project's effectiveness will be meta-evaluated according to its anticipated purposes. These are: (1) improved accountability; (2) increased sense of teamwork and commitment amongst professional staff; (3) improved decision-making; and (4) gain of public support for education.

1. Accountability

Three indicators of improved accountability are (1) the success of the evaluation model in information sharing and productive discussion of educational issues amongst those participating in the evaluation; (2) the efforts made by districts to distribute evaluation results to those not participating in the evaluation; and (3) the degree to which parents, teachers, administrators and/or trustees not involved in the evaluation subsequently engage in meaningful discussion.

Significant efforts were made by the districts to collect and discuss evaluation information. Public surveys, Stakeholder Team analysis, the production of detailed evaluation reports, and a number of follow-up activities demonstrate this quality. Stakeholder members described the process as "open", "educative", "rational" "interesting" and "meaningful" in terms of information sharing within the Team itself (see Appendix F5). With regard to this measure the IPS succeeded in its intent.

A second measure of the evaluation's success in accountability are the efforts made by the district to distribute the evaluation information. These efforts were moderately successful. The greatest effort was made in the Interior district; more individuals shared in the responsibility of promoting more events. In the Inlet district almost the whole burden for this task fell on the shoulders of the superintendent. He also developed the most creative accountability exercises; these were created almost out of desperation--because no one else stepped forward to assist, and because he himself was so committed to the concept. In the South Coast District the fewest accountability efforts were made. Yet the decision to produce a public newspaper insert ensured that the evaluation information was more widely distributed there than in either of the other two districts.

Factors that seemed to influence the information dissemination effect were: (1) the form of report decided upon; (2) the number of leaders who participated in accountability measures; (3) the initial efforts made to raise the profile of the project; and (4) the follow-up strategies devised by the district leadership. Efforts can be described as moderately effective. Subsequent attempts might be much more productive in impact because of the lessons learned in the first such venture.

The project was less successful in the third measure of accountability: engendering discussion of educational issues amongst those not involved in the evaluation. There is little evidence to show, for example, that the public in the South Coast District entered into any substantive discussion of educational issues as a result of the newspaper insert. No letters to the editor resulted. Stakeholder Team members indicated the level of use of the evaluation reports by both teachers and the public as "little". Some discussion of the issues took place informally, both at the district office and in schools, but the lack of a structured response mechanism meant that informal discussion was the only vehicle for debate. Little spontaneous reaction was forthcoming.

In the Inlet and Interior districts public and professional debate happened when it was structured to happen. The evaluation reports have the potential for creating meaningful debate on education, but processes and procedures must be employed to facilitate discussion. Efforts by the administrative team in the Interior district and by the superintendent in the Inlet district to accomplish this objective were termed "extensive" by most stakeholder members.

In summary, then, use of the evaluation model in the three districts has resulted in partial success with respect to the goal of improved accountability. Within the stakeholder group and those taking part in the evaluation, improved accountability did result. The effect in terms of spreading information and creating discussion amongst non-evaluators was minimal. This is so even with respect to the professional staff in the district⁵. There is potential for improved accountability. But success is a function of the administrative commitment and skill of the district leaders, a commitment that

⁵ Witness the fact that in two districts (Inlet and Interior) discussions took place at the staff level when encouraged by the central district staff, and did not, in the South Coast district, without such encouragement.

must extend beyond a single event. Systematic use of the IPS over at least three years could provide a more reliable determination as to whether its accountability goal has been achieved.

2. Teamwork and Commitment

A pattern of results similar to that in the accountability section was discovered when gauging whether the IPS created an increased sense of teamwork and commitment amongst professional staff in the districts. Much evidence exists to support an affirmative answer for Stakeholder Team members, but not so for the remainder of the district's employees, particularly teachers.

Two methods were employed to measure the System's success with regard to stakeholder teamwork and commitment. The first was through semantic differential analysis of words chosen to describe their role in the evaluation process. Each interviewee was first asked to place 36 word cards into one of three categories: describes role well, describes role somewhat, and describes role not at all. Then they were asked to choose three words that best described their role (the list of words chosen to describe stakeholder roles as individuals and as Teams is in Appendix F7). Secondly, interview responses were content analysed for evidence of teamwork and commitment.

Analysis of the word selections reveals the following trends. With regard to the three most popular words to describe the Team's role, Inlet stakeholders chose the terms cooperating, positive, involved, and partnership most often. South Coast representatives chose the same four words; and Interior stakeholders described their efforts as a positive, active, partnership. For the group of 21 stakeholders the words most often selected as the three most desirable descriptors of individual roles were cooperating, involved, and partnership; and cooperating, positive, and partnership were selected for the Team roles. A major difference between the three districts was the choice of the term "compliant" on two occasions in the South Coast district. This term reflected the belief of two team members who felt the process and the facilitator was too directive at times.

The words most frequently chosen to describe individual perceptions of roles are shown in Table 7.1: (see Appendix F7 for the complete list and frequency of choice). These same descriptors were identical for each district

and the districts combined, save one: In the South Coast district, the word managing was chosen rather than strong. Other comments such as: "no one did it all," it was "a shared responsibility," "there was no one who didn't take part" and they "really participated--every one of them," give credence to the assertion that the process developed teamwork and commitment.

The role descriptors contained within Table 7.1 bespeak of roles invested with a sense of productive, positive accomplishment--or for want of a better term, "empowerment." Jean Holmes commented, for example that:

...people listened to each other...I was really impressed with how people spoke up, even the parent representative...in a group of professionals a person like that could be intimidated.

Table 7.1: The 12 words most frequently chosen as describing individual stakeholder roles in the three district evaluations: by district and overall (words are arranged in descending order of importance; words with similar ratings are grouped together).

<u>Inlet District</u>	<u>South Coast District</u>	<u>Interior District</u>	<u>Overall</u>
cooperating	cooperating	cooperating	cooperating
positive	involved	involved	
	partnership	participant	involved
involved	participant		participant
participant	trusted	collaborator	
collaborator		partnership	positive
active	positive		collaborator
	active	active	partnership
partnership	collaborator	positive	
trusted			active
	<u>managing</u>	trusted	
knowledgeable		important	trusted
	decisionmaker		
important		knowledgeable	important
<u>strong</u>	knowledgeable	<u>strong</u>	
	important	decisionmaker	knowledgeable
decisionmaker			
			<u>strong</u>
			decisionmaker

This same parent described her own role as "cooperating, trusted, and powerful," and stated that there was "an underlying feeling that this was

important stuff, that it was empowering to be part of it." All but four relegated the term "powerless" and all but one the term "weak" to the category "does not describe role well." One Stakeholder Team member stated that "we were all made to feel we were important, and...(that we)...had a stake in it (the evaluation)."⁶ A second commented that the process "actually increased my level of expertise in terms of dealing with other people, as a result of that involvement,"⁷ and a third that he felt "powerful, opinionated...when I use words...(like that)...I'm trying to emphasize that I was not minimized, or my role as a trustee minimized."⁸ The groups were task-oriented and completed all responsibilities within given timelines; they "made decisions."

Contributing to the sense of "empowerment" was a feeling that the evaluation was done in an ethical manner. In two of the three districts there was strong stakeholder endorsement of the terms "honest", "just", "has integrity", and "open" as descriptors of the process; in the third district (South Coast) stakeholder reaction was less supportive but still significantly skewed to the positive side (see Appendix F8). Stakeholders in the Inlet and Interior evaluations felt that the project's processes and procedures dignified constituent opinions, that politics were conducted responsibly, and there was ample freedom within guidelines (see Appendix F9). However, South Coast stakeholders did not share the same views: three of the five felt the process was too rule-bound. This is likely due to the frustration they felt towards the Ministry questionnaire, and the timeline, which they also felt was "imposed" upon them.

A second ethical aspect relates to the issues of trust and control. No-one felt threatened; on the other hand, not all agreed that there was minimal use of executive authority (see Appendix F9). In the Inlet evaluation, for example:

...there was a lot more control by the superintendent than I would like to have seen...it is very difficult to be honest and open about it

⁶ Sam Gordon, Inlet School District.

⁷ Bill Blair, Interior School District.

⁸ Colin Scott, Interior School District.

all when you've got the head of it all...and they take it all very personally.⁹

In the South Coast evaluation it was the facilitator and the evaluation model that seemed somewhat authoritative. In the case of the Interior district the superintendent's attempt to run a meeting met with some resistance. One respondent who rated executive authority as too high was the teacher representative: this is the only indication that executive authority may not have been properly applied.¹⁰

A third ethical factor contributing to empowerment was "authority commensurate with responsibility". At issue here is whether a provincially sponsored evaluation process usurps local autonomy in decision-making. This was important to stakeholder members at the beginning of the exercise, and indications are that in the case of the Inlet and Interior district evaluations local autonomy was achieved. All Interior stakeholders, for example, answered in the affirmative when asked this question. And one Inlet stakeholder stated, "Right from the beginning the uniqueness of the local district was taken into account."¹¹ This was primarily due to granting the school district total control of all stages of the evaluation throughout the process.

Stakeholders in the South Coast district were not as positive about the ability of the evaluation to preserve local authority. Although they all felt the report reflected the uniqueness of the district, at least two were of the opinion that the process and procedures of the evaluation (i.e., the questionnaire, the timeline, and their inability to change them) hampered their ability to gather accurate and meaningful information. Whether this was truly a function of the process, or their ability to take advantage of the freedom granted them is immaterial: they did not feel the process provided for their district's uniqueness as much as it should.

⁹ Joan Holmes, Inlet District.

¹⁰ The director, Damon Korchuk, gave a 5 rating and the superintendent a 4 rating to this characteristic. However, the superintendent was of the opinion that this was the best rating; assuming the two extremes were undesirable.

¹¹ Don Gordon, Inlet District.

In contrast to the ability of the evaluation to create a sense of ethics, teamwork, and empowerment within the stakeholder team, there is little evidence that these qualities were transferred from stakeholder representatives to their interest groups. Teacher use of the reports was consistently described as minimal. Each teacher representative indicated that he was dissatisfied with both the groundwork and follow-up procedures. Indeed, the most common remark was that few people other than those involved in the evaluation even knew what was going on.

"Empowerment" influence seemed basically confined to the Stakeholder Teams themselves. It was expected that respect and teamwork that was created within a stakeholder committee would lead to greater cooperation between the district office and teachers. But this did not happen. In the Inlet case study, for example, teacher executives did little to assist with the use of the evaluation report. In the Interior district the teacher representative suggested that stakeholder groups should have input into planning, yet he took no initiatives to bring this about. And in the South Coast district, there was no official reaction from the Teachers' Association at all.

A possible explanation for the lack of transferred empowerment relates to the dynamics of teacher-management relations in B.C. A great deal of distrust still prevails. In the midst of these projects the BCTF actually issued a statement urging its members not to cooperate with Ministry accountability initiatives.¹² Although this political stance never became operational in any of the three evaluation projects, it is not surprising that voluntary teacher assistance with follow-up to this project--at least, initiated by themselves--has not been forthcoming.

In conclusion it must be stated that whereas the evaluation project succeeded in creating teamwork and commitment amongst stakeholder team members, it had little impact beyond the committee itself. Two contributing factors to this success appear to be its ethics and its success in fostering "ownership" of the results (achieved to varying degrees in the districts, as described in each case study). Until the sense of ownership can be

¹² Royal Commission Recommendation 7.18, in which districts are encouraged to develop accountability mechanisms in partnership with the Ministry, was listed as a recommendation "opposed" by the BCTF in the position paper entitled, Discussion Guide on the Royal Commission on Education, 1988.

transferred to all educators in the district, however, the project cannot be termed a success. This will not come about through a single evaluation event, nor without the process being built directly into the administrative fabric of the district.

3. Decision Making

Compared to its potential, the use of the IPS in the three districts had minimal impact on decision making. Only one of the three districts (Interior) actually built the information formally into its planning process and in this instance the process was already in place. In the instance of Inlet no formal planning structure was implemented; and in South Coast the evaluation information was informally referred to in the district's yearly goal setting process. No district committed itself to a long term planning cycle nor unequivocally to a repeat of the evaluation process, (although all three indicated informally that they will do so).

Stakeholder responses indicate that the evaluation model appears to have potential to contribute to the development of long term plans. It served to provide a focus for decision making, in that it provided a consistent knowledge base for members of the management team to make better decisions. It also helped to provide a clear vision of where the education system is and should go in the future, while considering immediate values needs and concerns as well as the more "long term" values of education. These were identified by stakeholder team members in two of the three districts as important achievements of the process (South Coast stakeholders questioned its abilities in these areas; see Appendix F10). In addition, stakeholder members saw the process as a practical tool for leadership, with only one person in 21 disagreeing. Damon Korchuk summed up the prevailing attitude with respect to the IPS's potential when he stated:

...I think we are only at stage one of possibly a three and maybe a four stage process now...I feel that there's more to come and I still can't see where the big picture is...

The evaluation process did provide an impetus for short term planning. The superintendent of Inlet supported this when he stated that the process

"does do a good job of establishing and communicating goals of a district...in long term planning we have obviously to get it together."

Stakeholder members in all districts agreed that long term planning would not happen unless two things took place. The first is that the evaluation and follow-up must become systematic. A second is to have, as Jean Holmes put it, "a structured process for follow-up...If its going to have use within the educational community or the public in the valley, then there has to be some kind of process." James Oord, the CIDTA representative in the Interior District agreed, stressing effective planning is possible "...if leadership is looked at in a cooperative sense."

A second measure of the evaluation's effectiveness in decision making relates to its impact on students. Since it is designed as an administrative tool, it is expected to have an indirect rather than a direct effect in this domain. However, if the advantages it provides for planning, for information flow, etc. do not ultimately improve student learning, then the evaluation cannot be termed a success. Stakeholder members did not see the process as having a meaningful impact in this regard (see Appendix F11). However, the IPS is intended to have its impact on students over a much longer term than is represented by this study. A longitudinal study using the measures of student improvement identified through key indicators would be a more effective measure of the evaluation's success. And since the IPS was designed to be used systematically, only when it has been systematically employed would a true measure of its decision making effectiveness be known.

There appears to be three major reasons for the failure to apply the evaluations to decision making. First, district leaders had difficulty envisioning the nature and structure of the potential planning process. None of them appeared willing to consider any administrative reconstruction to accomplish this; indeed, it is not sure that the possibility was ever really entertained. Those districts that had existing planning processes--such as the Interior District--simply adapted the evaluation information to suit it.

A second reason--and probably a consequence of the first--is that the goal of "public relations" seemed to dominate decision making throughout the evaluation project. Even when stakeholders saw the public relations value of "selling" a commitment to long range planning, they seemed either unwilling or unable to go beyond the production of an attractive annual report. In no

district was the concept of a planning report seriously considered. Public reaction to information always seemed in the forefront of people's minds. For this reason (in at least two of the three districts) truly controversial information (i.e., in Inlet, student reaction to school administrators; in South Coast, the issue of student transfer amongst secondary schools) was either not collected or omitted from the report. A decision to deal with the issue internally, or to ignore it altogether was a function of overemphasis on the public relations aspect of the evaluation process. This is further evidence of the "pathology of superficiality" described by Hodgkinson (1978); in which administrators give the appearance of meaningful endeavor, but truly only give superficial attention to a process.

A third reason--and one that will be discussed further later (Chapter VIII)--is that long term planning based on information requires a particular blend of administrative skills that probably few educational administrators possess. This appears to be a function of two qualities: (1) the ability to develop a long term perspective; and (2) training and experience as related to ethical and technical expertise in the area of information retrieval and management. The IPS model was a powerful tool in information retrieval--but skills in conversion of this information, fairly and meaningfully, into a plan of action--were not skills demonstrated by the educational leaders in any but the Interior district, and there only in a pure public relations fashion.

Despite its failure to spur the development of long term planning, the evaluation model certainly succeeded in identifying immediate district needs and bringing them to the awareness of decision makers. In the South Coast and Interior Districts existing goal-setting processes took advantage of the evaluation's information, probably with more effect in the latter district (in that it influenced budget decisions). In addition, there were indications in each case study whereby stakeholders became more knowledgeable about the district and pursued actions to improve it. This was the case both with information contained within the formal reports, and from informal knowledge gained through the process--information relating to people's style, district perspective and constituent opinion gleaned from the questionnaire. This demonstrates an educative use for the evaluation project (see Appendix F5). But without a holistic plan of action in which priorities are established, responsibilities clarified, and timelines set, the press of immediate events will

soon muddy these perspectives, and blunt the evaluation's effectiveness. The potential for political pressure and emotion to dictate action becomes heightened. Without an overall plan, the evaluation has become a singular event, rather than a stage in a systematic renewal process.

4. Public Support

There is mixed evidence that the Inlet study succeeded in accomplishing the objective of gaining public support for education. Stakeholder team members did think that the evaluation report provides a consistent knowledge base for the general public so they could gain a better understanding of the priorities and achievements of the local education system, but they did not agree that it would lead to a higher degree of shared commitment and purpose amongst the parents and the general public (see Appendix F12). This could be due to the fact that the evaluation was new, or that the report was not as readable as it might have been, or that follow-up procedures were not well laid out in advance.

The community response in Inlet did not indicate a groundswell of increased support. Employers did not show an interest in discussing the evaluation report with the superintendent. Reactions to the report (in terms of written reactions sent to the school board) were almost nonexistent. And employers did not respond in any but the most passive way to the overtures made in their direction. This was the superintendent's biggest disappointment: "I can't seem to arouse as much school or community reaction as I want...I'm having a real trouble finding people who care." South Coast administrators employed no stratagems to create community commitment or response, and therefore none was forthcoming. Interior administrators, on the other hand, carried out a well-defined plan of action that certainly could not avoid creating some public awareness as to the state of conditions in the school district. The video played a major role in this strategy. A follow-up study would be required to measure the ultimate impact of this upon people's attitudes.

It should also be pointed out that the superintendent of the Interior District was only one of the three to formally employ others (i.e., district staff, parents, and administrators) as disseminators of the evaluation's information. By asking them to discuss the report at public meetings; by providing them with a

video; and by distributing the booklets to staff and the community, he delegated the responsibility of follow-up and assisted its implementation. In contrast, all activities in the Inlet district were initiatives by the superintendent--even the novel and interesting debate in the newspaper. With only one individual trying to raise awareness the effect is going to be minimal.

In summary, it should be clear that motivation and gain of public support is a function of two things: (1) whether or not a follow-up strategy to address this need exists; and (2) whether sufficient creativity, resources, personnel, and expertise is available so as to be able to structure a comprehensive district dialogue (even with both of these factors in place, the administrator must still ensure that the discussions take place ethically, and are not used purely for political manipulation: i.e., to control people's opinions, rather than sincerely seek out improvement suggestions). Evidence exists to show that when a response process was structured--like at the junior and senior secondary school parent meetings in Inlet, or the Chamber of Commerce in Interior--people reacted positively. Individuals at the junior high meeting, for example, did indicate in their summary that they were more willing to participate in educational improvement than they were before the evaluation. These pockets of support will likely remain isolated until attitudes, behaviors, and habits are changed through systematic use of the IPS; or until an overall planning process, involving the community, is built-in to the structure.

Summary

There appeared to be six factors delimiting success with regards to the effectiveness of the IPS in achieving its goals of accountability, improved professional teamwork and commitment, better decision making and the gain of public support for education. These are: (1) the political/administrative importance the project assumes in the district; (2) the existing patterns of leadership and style in the district; (3) the personal ability and style of each actor in the process; (4) the interaction between these qualities and the styles and skills required by the evaluation project; (5) the follow-up plan developed in the district; and (6) resources and situational elements found in the context of each district. A closer look at the limitations to evaluation use, and its role in administration will be examined at greater length in the next chapter.

CHAPTER VIII

General Discussion and Commentary on Evaluation and Leader-substitute Theory

A major goal of the study was to observe and gain further understanding as to why formative evaluation has not been used effectively for improvement of educational organizations. Analysis of the three case studies provides some insights into this theme. In addition, the dynamics of administration revealed throughout the study provide opportunities for insights into the nature of administration and the leader-substitute construct of leadership.

Evaluation

Discussion will take place in two stages: (1) Observed Limitations to Evaluation Effectiveness; and (2) Insights into Formative Evaluation in Educational Administration.

1. Observed Limitations

Evidence exists to show that some of the traditional reasons given in the literature for the failure of evaluation were operational in the three evaluation projects. In some cases the processes and procedures overcame them; in others, the effectiveness of the process was curbed by them. A number of other possible explanations for the limited success of the evaluation also suggest themselves.

a. Traditional limitations. The literature suggests two main categories of reasons as to why evaluation effectiveness is limited. The first category is entitled "political reasons", and the second "technical reasons".

Political reasons refer to antipathy towards evaluation because it might be used to manipulate people or groups. Evaluation is threatening because it can "point the finger" at individuals for the poor performance of a program. For example Inlet superintendent Bob Ney "was a little defensive at times...it's understandable...its the superintendent's district, he'll take the flak if everything isn't ok"¹. The director in South Coast experienced the same

¹ Eva Moran, Inlet School District.

circumstance with individual teachers: some of the information (exam scores, for example) identified individuals and they became quite defensive and he had to spend a good deal of time settling them down.

Groups perceive evaluation as potentially open to abuse in that information may be used to further the political purposes of one group at the expense of another. Fears by the three teacher's associations that the questionnaire process would be used to further the political interests of the school board was a factor in gaining initial acceptance of the project. Frank Ford, of Inlet, for example stated that "we have a good relationship with the teachers, and I was afraid we would jeopardize it." And Bob Ney said, about the stakeholder meetings,

...if at any time...the discussion swung to giving information that served a political purpose, then...(cooperation)...would have collapsed.

In South Coast stakeholder members were only partly satisfied by being given an equal role in interpretation of an 'imposed' survey's findings. Yet, teachers in the Interior district expressed satisfaction because they were involved in all aspects of survey development. While political fears were assuaged to some extent by the stakeholder collaborative process, they resurfaced during evaluation follow-up, and are likely a reason why teacher participation was so limited. Sally Page stated, for example, that "I've had some indications from teachers that their concerns that the report would be used in a negative way is perhaps coming true." Would this have been the case if a collaborative planning process followed the evaluation? And to what extent would this be attributable to a newly unionized Teachers' organization that is dedicated to subsystem values at odds with those of the organization as a whole?

Technical reasons for evaluation limitation relate to methodology, resource allocation, and expertise. Limitations with respect to all but the latter have been discussed at length (i.e., timing, methods, etc.) in the case studies and meta-evaluation chapter. However, a few words need to be said concerning the topic of expertise. Administrators would require specific training or professional assistance in the technical aspects of evaluation--surveying, data analysis, interpretation, report-writing, consensus building, etc.--before the

IPS could be used regularly for educational improvement. This will be discussed at length later.

b. Other limitations. Experiences in the three case studies suggest other, less traditional reasons for the limited effectiveness of the evaluations. One such reason is the lack of awareness as to the true complexity of the act of evaluation itself. Earlier in the paper it was described as a three stage process, in which an object is measured, compared to an ideal, and then discrepancy decisions as to merit and worth are made. When an object is complex, like a school system, there are ever-increasing levels of value generalizations created as one moves through these three stages. Each level is composed of layer upon layer of value judgment, and progress through them is dependent on rational generalization from one level to the next. Unless the process is carefully structured and managed, individuals who represent different organizational and personal interests are prone to make large leaps of judgement, leading them off in different directions. To illustrate this complexity, a list of the value judgments required in just the first stage of the Inlet evaluation are shown in Appendix G1.

The limitation due to complexity is exacerbated by traditional role responsibilities in school districts. Traditional bureaucratic evaluation encourages individuals to conduct evaluation at their level of responsibility only. This leads, deliberately, to a reliance on personal judgment, knowledge and intuition--rather than organizational, or shared, judgment, knowledge and intuition. In the former case evaluation can quickly be identified as self-serving and "unfair"--a form of control in a hierarchical structure, rather than a natural process of organizational growth. Without any formalized structure to share, analyze, and synthesize evaluative decisions the organization has to depend on individual rationality, knowledge, and ability for problem-solving. The many layers of value judgment inherent in organizational evaluation, if not managed carefully, rationally, and holistically will likely act as a major limitation on its acceptance as a valued process in organizational development. In this study the districts that appeared to rely most on bureaucratic management techniques were the ones that realized the least return from the project.

Another limitation to evaluation effectiveness was the need for organized follow-up. The extent to which this was planned seemed to influence its

success. Evaluation is often perceived as the last activity in a program, or plan. As such its power to provide the knowledge, vision, and focus required to face future challenges is often not automatically translated into the practice of leadership. An evaluation process in and of itself can provide direction and focus, but if this direction and focus is not channeled through a commitment to an organized decision making process, the impetus for long term decision making soon is overcome by immediate events. Like follow-through on a golf swing, or organizing and planning a diet, evaluation will quickly go awry if effective follow-up procedures are not developed in advance. How evaluation information will be used, by whom, for what purposes, and for what term, should be decided prior to the evaluation, and adhered to regardless of circumstantial pressures.

A fourth observed limitation to evaluation effectiveness was the level of trust that existed between individuals and groups in the school districts. The intent of formative evaluation is to pinpoint directions for improvement. Because improvement tends to suggest weakness, or a lack of competence on behalf of someone or some program, individuals or groups must trust the motives of the other people they are working with. Distrust of Ministry intents (South Coast); of how each group within the districts would use the evaluation (Inlet); and of the motives of the superintendent (Inlet & Interior); led to a great deal of apprehension at the beginning of each case study. The stakeholder interactions seemed to overcome this limitation, at least as far as internal trust was concerned, only to have it resurface when the committee ceased to meet and follow-up was required. This was particularly true in South Coast and Inlet, and less so in the Interior district.

For reasons of trust it is also fundamental that organizational evaluations--such as the Inlet study--be clearly separated in function from personnel evaluations. One should conduct an organizational evaluation with the idea that roles are being examined, and not individual competence. It should also be noted that if personnel evaluation is practiced poorly in the district, or is hamstrung in its effectiveness by legal constraints, the temptation to use system evaluation as a surrogate is great. Succumbing to such temptation will destroy trust immediately, and curtail the evaluation's effectiveness.

A fifth possible limitation relates to the issue of people's motivations. During each evaluation, and particularly at the time when the survey results

were being analyzed, one got the sense that there was a tremendous sense of relief that the statistics did not seem to point to major problems in the district. This satisfaction with their own achievement seemed to diminish the thrust for improvement. Individuals seemed impelled to make summative judgments as to the district's worth, and if that judgment was "OK" then thrust for improvement in specific areas of weakness was somewhat blunted. A second aspect relates to the concepts of "approval" and "blame". Teachers in the Inlet district, for example, resented the public relations purpose of the evaluation because it implied blame for education's poor image. The desire for approval seemed to drive Stakeholder Teams away from the planning goal in the direction of the public relations goal; strengths were sought out and highlighted much more than weaknesses. One got the feeling that information would be resisted if blame was apportioned, but would be accepted if approval was forthcoming. Yet, when that approval was manifest--either through inferences from the data, from each other, or from a sense of pride in the project itself--people were satisfied, and ready to move on to new things. It was almost as if the evaluation was exciting and interesting, but follow-through was dull and boring; yet it is the follow-through that creates the true formative nature of the evaluation process. All the more reason to plan it beforehand; to plan it so prior commitment can stimulate the will to pursue long term goals rather than short term ones.

The sixth possible limitation suggested in the three studies has to do with the concept of "resistance to change." Evaluation has the power to suggest needed changes, yet it is acknowledged that it is human nature to resist change. First evaluations--such as the case studies--suggest major change, because cyclical, systematic evaluation has not taken place so as to allow the organization to adjust naturally to new challenges. It could very well be that the changes suggested by this evaluation--to leadership in the schools in the Inlet District, for example--were so fundamental that they could not be accomplished without major alterations in district practice. If this was the case, then "event" evaluations that force people to make changes that strongly challenge existing habits or ideologies or existing power relationships or that challenge one's own role identity, will be resisted, or allowed to fade away. One must feel they have the ability to conduct the changes required; and one must have the will to do so.

A seventh limitation to evaluation effectiveness is suggested by the apparent need of all groups to "internalize" or own the process. Without the administrators onside, without the public being aware as to what were the purposes and processes of the evaluation, and without teachers being committed to it the evaluation's potential was limited. This appears to suggest that somehow individual values and organizational values must be reconciled during the process of evaluation. Whether in fact strategies to overcome this limitation will also overcome "natural resistance to change" remains to be seen in future studies.

An eighth possible limitation to evaluation effectiveness is to some extent a technical one, to some extent a fundamentally philosophical one. Many, many hours were spent in each district trying to decide how to present information accurately. An example was the debate on suspension rates in the South Coast district. Should suspension rates be represented as "number of suspensions per year," as "number and type of suspensions per year," as "number of days lost to suspension as a percent of total attendance days," or as "percent of students not suspended for a school year"? Should the information be presented over time, in a table or in a graph, or in prose? Should the graph be small or large, and should it be accompanied with pictures of students behaving or misbehaving? What other information is necessary to present this data in a fair light? These and many other related questions soon led one to an awareness that the value judgments inherent in information presentation almost belie the meaning of the word, "accurate". Where is the line between accurate presentation of information and public relations presentation of information? In a society in which information is plentiful, the real challenge in evaluation is how to manage information in an ethical and meaningful way. What is the right information, the right amount of information, and how should it be presented? Without the ability to answer these questions in a practical as well as philosophical sense, evaluation's effectiveness is severely limited.

These "observed" limitations to evaluation effectiveness are specific to these case studies. However, an understanding of these limitations is facilitated when one examines formative evaluation in terms of the factors mentioned at the conclusion of the previous chapter--factors relating to administrative style, expertise, skill, and district context.

2. Evaluation and Administration

An awareness that grew throughout the evaluation project and one that crystallized into almost a conviction near the end was that three basic assumptions derived from the literature about formative evaluation should be challenged. They are: (1) that by developing an evaluation model to complete the administrative learning cycle (such as in the IPS), formative evaluation has been implemented; (2) that educational administrator experience with evaluation gathered at the classroom level provides a good base for the use of evaluation at the organizational level; and (3) that theoretical commitment to formative evaluation would translate into a practical commitment when models to implement the theory are employed.

The researcher assumed at the beginning of the three case studies that once the values framework (i.e., district vision) and key indicators of performance had been established through the use of the IPS model, district administrators would automatically use the specific information in a logical follow-up process dedicated to district improvement. Each IPS implementation became a one time only event; moreover, it quickly became apparent that one isolated implementation does not necessarily lead to habitual use. Formative evaluation means utilizing the evaluation processes systematically, in an ongoing way, to inform all levels of the administrative learning loop through a higher level of judgment (see Figure 8.1). In other words, it is the follow-up that is important in formative evaluation, not the event itself. All the episode does is model how evaluative decisions can be made, and identify what information is key to decision making with respect to the goals of the institution. Follow-up must build them into the day-to-day operations of the district. The follow-up must try to maintain the logic of the process so as to sustain productive dialogue amongst stakeholders; it should seek constant updates of fluctuations in indicators; it should inform decision making through consensual judgment as to what those fluctuations mean in terms of district performance. By so doing evaluation would become the key to development of a holistic, learning, active-reactive-adaptive organizational pattern. This stage was lacking in each district, and the main reason was that it was lacking in the original model.

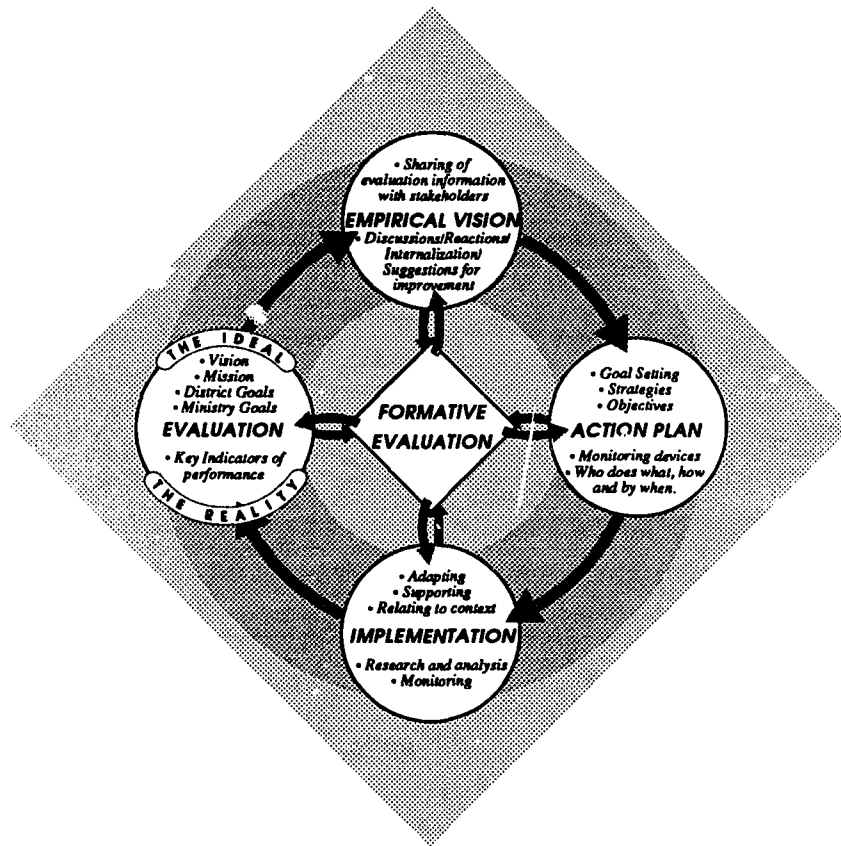


Figure 8.1: Formative evaluation and the learning loop for organizational renewal.

The second assumption that needs to be challenged relates to the premise that "evaluation is a well accepted element in the management of education" (McPherson, Crowson & Pitner, 1986, p. 84). These authors argue that "the education profession is no stranger to the use of evaluation" because "testing and measurement techniques have been elements in training programs for teachers"; and students, teachers, and principals themselves are constantly subjected to personnel evaluations. It is, however, a significant leap from these arguments to the assumption that formative evaluation is a well accepted element in the management of education; indeed, the examples given are all presented in the sense of summative evaluation for the purpose of personnel control, not personal growth.

A corollary to this assumption is that because educational administrators have gathered data, interpreted data, and communicated value judgments in the process of teaching, then they are able to do the same with data at the

district level. Two premises underlie the corollary: the first is that teachers are skilled in data manipulation; and the second is that they are motivated, rather than required, to do so. Both premises are false. Some educators are effective student evaluators, some are well-trained in the development of evaluation models and strategies, and some are proficient in their use. However, this is far from being the educators' true strength. Most educators possess very high levels of people management skills, the very skills that are able to mask their lack of expertise in the area of effective data collection, evaluation, and reporting. Indeed, the source of much of the public concern about education relates to an inability to provide evidence as to what students can do, and what the system has done to influence them. Educators rarely receive training in evaluation; grading practices within schools are often arbitrary, norm-referenced, and poorly planned. Proper testing techniques have been haphazardly learned, and the prime role of evaluation is control, rather than formative growth for the student. For example, how does the most powerful symbol of student achievement, a letter grade, indicate directions for improvement, particularly if it is often a function of innate ability? Many teachers resent having to justify evaluations through effective data keeping and the presentation of information; they prefer to rely on so-called professional judgment-- even if that judgment cannot be understood by either students or parents.

These comments are occasioned because educators in the three case studies found it problematic to: (1) understand and apply formative evaluation strategies at the organizational level; (2) manage, analyze, and interpret meaningful information with regards to what students achieve in the educational system; and (3) report what is meaningful in a language and manner that creates understanding of the purposes and achievements of the education system. These skills seem more well-developed amongst newspaper columnists and TV journalists than in the teaching profession; if "management of meaning" is important to educational administrators, then much more needs to be done by way of education in this domain. Indeed, the low priority given evaluation may be because administrators recognize this weakness and avoid the responsibility as much as possible.

What evidence was provided by the three case studies that lead to the conclusions about the understanding and skill level of administrators with

respect to formative evaluation? First there was the pattern of response from senior administrators--and others--who made the general comment that they had a hard time understanding the "vision" of the evaluation project. Statements like, "We didn't really understand it at the beginning,"² "It is difficult to see the end,"³ and "I feel that there's more to come and I still can't see where the big picture is..."⁴ attest to this assertion. This, in spite of the fact that everyone embarking on the project knew that the main purpose of the project was improvement of the educational system through formative evaluation.

A second piece of evidence was the unwillingness of the administrators to truly differentiate the true function of formative evaluation--system improvement through long range planning--from the function of public relations. In all three districts public relations, or "positive" accountability became the prime focus of all of the follow-up activities. In the instance of South Coast, for example, the Director, John Lee, was unable to divorce the content of an internal planning report from the content of a public report; and in the Interior and Inlet districts, a decision to produce a public report took precedence over the management one.

Third, although most administrators articulated long range planning as a fundamental purpose for doing the evaluation, constructing the bridge between the evaluation information and plans for improvement seemed very, very difficult. Not only difficult to do, difficult to understand. The only district in which a real sense of cooperative long term planning seemed to be a possibility was the Interior District, in which cooperative short term planning mechanisms were already in place.

The inability to construct a bridge between evaluation and planning appeared to have two causes. The first seemed to be a lack of true understanding of what formative evaluation means. A second may be the result of the first; it appeared that administrators did not know how, even if

² Colin Scott, Interior School District.

³ John Lee, South Coast School District.

⁴ Damon Korchuk, Interior School district.

they understood the formative evaluation principles, to put them into some concrete plan of action.

District formative evaluation is a two stage process, with the real challenge at the second stage. First, there is the construction of a complete learning loop in which evaluation is an ongoing partner in normal decision making (see Figure 8.1). Second, it requires commitment to clearly defined follow-up procedures, consisting of monitoring and formative evaluation. Formative evaluation drives the learning cycle of activities--evaluation, vision setting, planning, and implementation--forward in an active-reactive-adaptive way (see Figure 8.2). And, it requires a commitment to change: a determination to adjust the organization to new and challenging demands placed on it, through the creation of an institutionalized formative evaluation process. Although agreed to in principle before the evaluation project, the case studies gave varying testimony to the ability of the IPS to accomplish the above requirements.

It should also be emphasized that commitment in principle and commitment in action are two different things. The above discussion implies limitations due to knowledge, skill, or technical expertise. The inability may in fact have been in the area of determination, desire, and ethics; or, for want of a better term, "will". Administrators may have understood formative evaluation, may have in fact known how they could have constructed the bridge to planning, but may not have had the fortitude, perseverance, or ethical skill needed to institute these measures while having to, at the same time, deal with other priority issues. It is for this reason that an understanding of the cyclical and structural nature of the formative evaluation procedure needs to be understood in its totality at the beginning, and that the key to development of such a process is to build evaluation methodologies and administrative practices used in the event into the follow-up. Timelines need to be established, teamwork structures developed, and time and resource commitments openly designated to the project. In this manner commitment is engendered, for these pre-project decisions bind the administrator to their successful completion.⁵ This same lack of will caused districts to move away

⁵ There are two models of long term planning processes that are currently being used by school districts in B.C. that exemplify this need: they are found in School District #62 (Saanich); and School District #43 (Coquitlam).

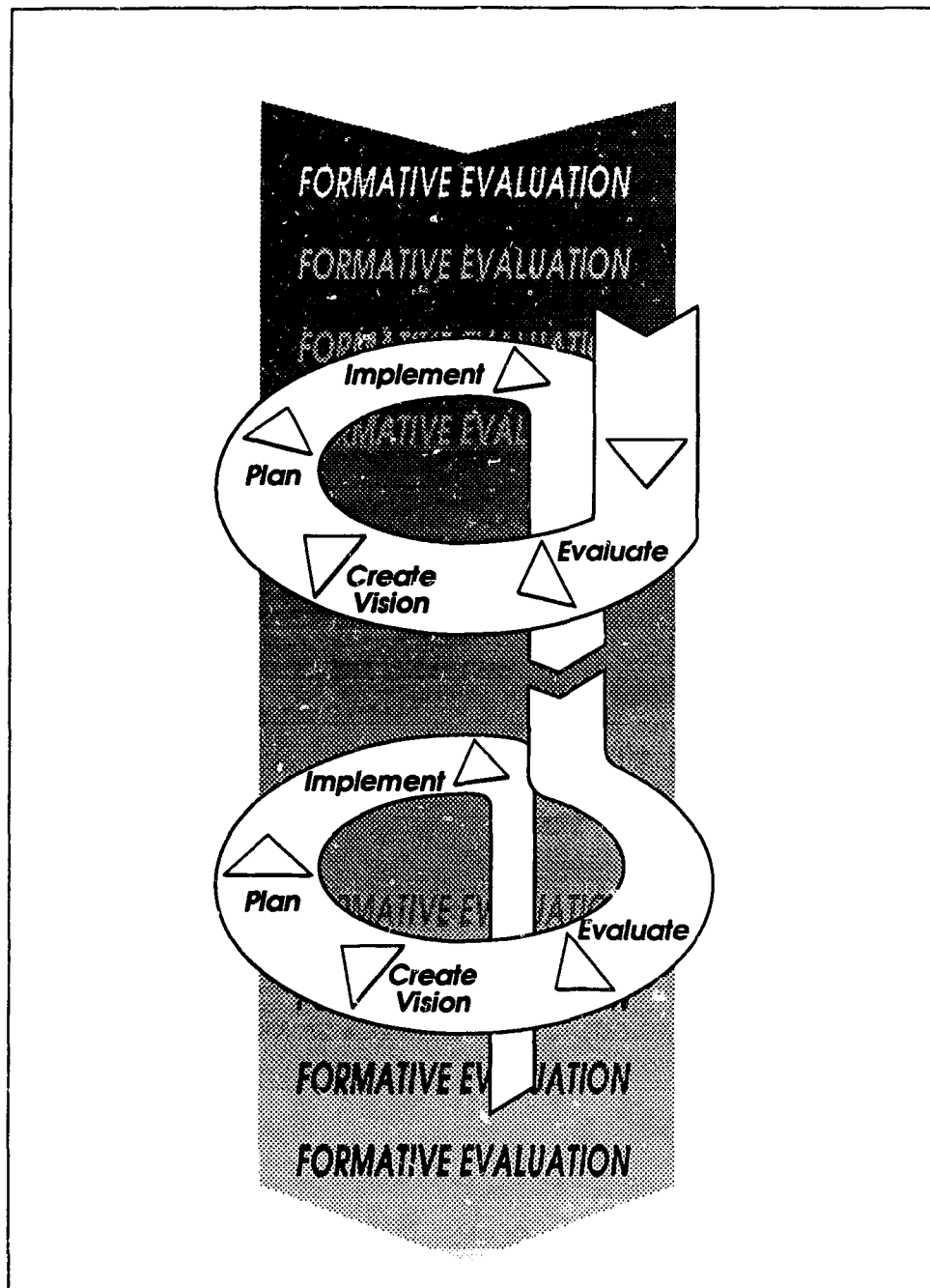


Figure 8.2: Formative evaluation as the force moving the organizational learning cycle forward in an active-adaptive-reactive way (adapted from Kemmis & McTaggart, 1988, p. 11)

from the planning goal--i.e., a commitment to the goal of growth and improvement--towards the public relations goal. It is partially responsible, too, for the lack of responsiveness to the superintendent's follow-up activities in the Inlet District; he did not design mechanisms to truly require follow-up. Joan Holmes emphasized this, stating:

...we ended up patting ourselves on the back too much...if we really are going to get any better we have to find out our weaknesses, be honest about them, and go forward from there...

A further limiting factor to formative evaluation effectiveness was mentioned earlier and needs further explanation. This was the concept of "ownership". It is related to the concepts of vision and will, because ownership means buying, or committing oneself to a plan of action due to a value congruency between one's personal interests and those of a project, or organization. The act of buying implies commitment; value congruency implies understanding and knowledge of the value purposes of the exercise. Ownership was repeatedly mentioned as a source of empowerment, and had much to do with attitudes to the questionnaire, to key indicators, to data choice and interpretation, and to evaluation reports and follow-up activities.

Ownership strategies were a combination of evaluation principles and administrative skill. Without the right to develop the questionnaire, to develop local key indicators of performance, to interpret the significance of the measure and decide how to report it, stakeholders would not "own" the result, would not view the evaluation as ethical, and would not commit themselves to the follow-up required to improve what needs to be improved. To create this ownership the IPS model evolved to emphasize effective groundwork, and to "sell" the formative evaluation vision. Multiperspectives gained through subgroup sampling was another ownership strategy. The whole purpose of building intuitive evaluation stages into the formal evaluation processes was to raise the value consciousness of organizational members, and to connect personal values to organizational values. The formation of a stakeholder team, and its role in key indicator development, and "group" data interpretation and report writing were strategies to resolve value conflicts and form value congruencies.

Evidence throughout all case studies indicated that when people expressed a sense of ownership they were more positive about the evaluation process or procedure, and more committed to follow-up procedures. Sandra Page, from Inlet, stated: "the whole area of accountability improves when it becomes a cooperative effort--i.e., not Gods from above pointing out strengths and weaknesses, and going away." She also commented that "currently teachers are cut out of planning...the school board is unable to create action plans for education; they are not competent...If direction comes from the district office, it is 'laid on' and harder to gain acceptance." Bill Blair agreed, remarking that evaluation has not traditionally involved stakeholders, and "in not involving stakeholders you cannot develop post-improvement activities."

Lack of ownership and the earlier contention that formative evaluation is not really understood by the teaching or administrative population combine to create most of the apprehension surrounding evaluation. People "fear" formative evaluation because they have difficulty differentiating it from summative evaluation, which is a form of bureaucratic control. The whole concept of role specialization and authority in a bureaucracy are such that ownership between the values of one role and the values of another are difficult to mesh. Dan Roebottom stated that "maybe someone won't agree with your approach"; and "evaluation is often based on the perceptions of the evaluators." And when those evaluators are superiors, whose organizational values in a "loosely coupled" organization do not relate to the realities of the situation you find yourself in, there is heightened apprehension and "evaluphobia". Bill Blair confirmed this when he stated that there "... haven't been appropriate evaluation devices that have allowed decision makers to progress and develop with a positive thrust." As Eva Moran put it, "most teachers want to be accountable, but are afraid as to how (sic). Evaluation that is false is worse than no evaluation at all." It is this very concept that is the basis of Patton's (1980) utilization-focused evaluation principles, the "organization" focus of the IPS, and the rationale for a stakeholder collaborative process to guide the evaluation.

It was stated earlier that people also fear formative evaluation because an implicit requirement is the requirement to change. Is it evaluation that is the source of threat, or is it the requirement to change? Note the comments of some of the stakeholders: "we see things we would like to change, but the

resources are just not available"⁶; "attitudes can be changed, but not easily"⁷; "Unless you really know what you want to evaluate its hard to avoid evaluating other things along with it...you get into things you don't want to know about"⁸; and

...I think staff feel very vulnerable to political decision making by boards...so why put the energy into it...(formative evaluation) ...when the probability of effective change is replaced by whimsical change? ⁹

The relationship between change and formative evaluation needs further examination, particularly as "ownership" is identified in change literature as fundamental to effective change strategy (Karr, 1988).

Ownership of evaluation results immediately creates ownership of the revealed problems. If formative evaluation is a powerful learning exercise, if it gathers information vital to decision making, if people are committed to action as a result--what happens if the evaluation points to actions that one feels one is unable or unwilling to take? Secondly, what if the procedures and processes established for making improvement decisions do not allow for input from those who now believe they "own" the problem, and now wish to be part of the solution's design? The latter circumstance certainly influenced the success of the follow-up activities in the Inlet district, and to a lesser degree in the Interior district.

Ownership of revealed problems becomes particularly problematic to administrators because it narrows the boundaries in which they are free to act. This problem is aggravated in public process. For administrative decisions are not made on public knowledge alone. Each decision is more likely the result of balancing public information with private. Colin Scott, the trustee from the Interior District, mentioned this fact when he discussed the role of the "informal" organization in evaluation and decision making. Ethical

⁶ Frank Ford, Inlet School District.

⁷ *ibid*

⁸ Ken Franks, South Coast School District.

⁹ Colin Scott, Interior School District.

action often requires the administrator to balance the message contained within explicit information with information one has learned informally about relationships, or values held by others. When taken in concert it may be that the ethics of the situation require the administrator to deny the action implicit in the public information in favor of one that is suggested by private information. This same quandary is extant in personnel evaluation. Denial of decisions implied by public information often forces the disclosure of the private information, and that in itself compounds the ethics of the issue. In this sense the accountability/public relations dialectic of the evaluation militates against its potential for effective decision making.

In evaluation models such as the IPS, whereby information becomes the currency upon which decision making is going to result, then one must be convinced that the quality of that information is such that it will not lead one into ethical conundrums. In each case study there was evidence to show that there were clear directions for improvement revealed through the evaluation process and/or report. A number of stakeholders commented that these required improvements were "obvious", made explicit by the information itself. The information became a goad to action, and made public, almost irresistibly so. Yet there was also ample evidence to show that information management created ethical concerns, in terms of determining its true meaning. This is one reason why teachers complained about exam scores and their use, why teachers worried about the wording of questionnaires, and why the superintendent of Interior wished to cancel the staff survey during the negotiations conflict. If the information is flawed, then it can create many more problems than it may solve.

This explanation reveals at least in part, why information management became so important in each district. People were constantly striving to express information in time trends; they sought better quality data, clearer meaning, and better data representation. To create quality data, cooperative value judgments, in a logical progression from specific to general, became the goal of the whole evaluation project. The cooperative judgments extended into data representation and meaning. Each graph, statistic, piece of prose, and quotation became a negotiated item. Everything but the most raw data became a function of subjectivity and opinion--albeit done through a process that tried to develop criteria processes to guide people through a rational train of

value judgments. Each person knew that the representation of that data would create demands for action--and they wanted to be sure that the demands so created would be ethical, meaningful, and manageable.

Information management was complicated by two other issues. The first is that no objective data or information was contained in any evaluation report. It was already reconstituted, reframed, and regeneralized in so many different ways so as to be chosen for its effect, rather than its pure essence. This was done to the degree that the essence was itself suspect. Second the complexity of group evaluation is such that without a process to guide them, evaluations quickly degenerate into pure intuitiveness, and ownership of a shared judgment is lost. Ownership requirements create a need for very sophisticated group management skills. Evaluation is therefore limited to the extent the individuals conducting the evaluation are able to employ them.

To summarize to this point: four causes have been postulated as to potential reasons why formative evaluation has had limited use by educational leaders in district development. The first was that administrators do not have a true understanding of formative evaluation. The second is that either they tend neither to have the data management nor the planning skills required to properly conduct formative evaluation. The third is that they lack the will to do so. A fourth is that weak as technical expertise in terms of data retrieval and representation may be, it outstrips to a great degree an understanding of the ethical use of such information. These limitations are worsened because experiences in this research suggest that "factual" information--traditionally seen as the source of "rational" action in any evaluation--may not even exist in school district evaluations.

Two other possible limitations to the effective use of formative evaluation by educational administrators were suggested by this study. The first has to do with psychological issues related to the individual. The second relates to whether organizations are best viewed as "holistic" or "mechanistic". A choice of viewpoint may influence administrative willingness to change.

The three cases in this study are examples of the attempt to apply "holistic" formative evaluation to school district organizations. Implicit in this endeavor is the belief that many of the fears associated with evaluation can be overcome if one evaluated the organization as a whole and not the individual organizational member. There was a deliberate attempt to overcome what

Scriven (1983b) calls "evalophobia"; to avoid the potential that blame would be visited upon a any individual for problems that a holistic view of organizations suggest are morphological. It is consistent with the holistic philosophy that it denies the role specificity and influence of bureaucracies, taking the point of view that the whole is greater (or lesser) than the sum of the parts; thus to evaluate so as to lay blame on a "part", or individual member, would have been an unrealistic form of evaluation.

However, even the holistic conception of organizations must come to grips with the psychological limitations of individuals. This study emphasizes a point made by Shapiro (1988); "The impact of trust on the evaluation literature has not received the attention it deserves." Evaluation highlights the need for "trust" and "commitment" to develop amongst individuals, and for "fear" to be removed as a consequence of evaluation. Fear is an emotion; to a great extent so are trust and commitment. These feelings are possessed by organizational members, including administrators. Within the stakeholder groups of each district these emotions were mixed and converted into a team thrust. The group thrust tended, as a whole, to overcome individual psychological limitations through the creation of trust; and once the group was disbanded, individual limitations once again asserted themselves. For example, in the South Coast evaluation it was noted that John Lee spent hours and hours proofing the district report, because "he might look bad," and because he wanted to ensure that "he would not be blamed for any mistakes in the report." He was not "willing to take any risks at all," and was "very pleased with those items that made the district look good." He even stated himself that "If I thought there were weaknesses...(in the report)...I wouldn't have had it published."

Evaluation effectiveness is affected by the emotional strength and moral fibre of the individuals taking part in it. So is change. These limitations can be somewhat overcome by trying to construct a group consciousness, as in the stakeholder teams. But does then the individual's self-esteem quickly become built into a district consciousness? Are there "organizational" emotions? Do these organizational emotions lead to the development of an organizational self-concept that, as it does in individuals, makes it extremely difficult to conduct a "rational" evaluation of self? Is "fear" of evaluation actually apprehension as to having to adjust one's own image of oneself by adjusting to

criticism? And once that criticism is forthcoming, and through internal analysis is rationalized, or explained away through the ownership processes of the evaluation, is then the goad to action blunted? Is this why follow-up in terms of improvement seems to get short shrift in favor of public relations--because it is easier to change the outward image when you know what people are looking for, than to really accept the need for internal change? What is the true psychological commitment to change in all of us? Norma Rivette stated, for example, that "people do not like to know areas of weakness" as individuals: why do we assume that as organizational members we are any more willing to know them and act on them through the use of evaluation?

The holistic conception of organizations--and the role of formative evaluation within them--also creates new images of the change process. In a mechanistic system, change is perceived as a complex but essentially linear extrapolation of cause and effect. In holistic organizations, change in one part implies the creation of a new and indeterminate whole. Change in the former instance is much easier to control than in the latter; progression into the future much easier to envision.

What evidence existed in the three cases to support the holistic concept? As pleased as most stakeholders were with the various evaluation reports, there was an awareness that they were more like mechanistic snapshots of districts, rather than holistic ones. A "print" evaluation report represents a view of the district only appropriate to a specific time and situation. Not only that, so much of the achievement of educational success is found in the relationships between individuals; how does one convey this through statistics, and show its dynamic aspect in the print medium? For this reason pictures were included--to emphasize the human relationship element in education. But only in the video report was the impact truly significant. Is this why traditional evaluation reports--inevitably two dimensional representations of a three dimensional enterprise--soon lose currency and interest¹⁰?

A holistic view of organizations also suggests why evaluation creates ethical challenges for the administrator. How does one understand the ethical

¹⁰ The video report "brought to life", through pictorial representation, the values of positive human relationships that are difficult to express in statistics. They represented those statistics in a visual way. Patton (1987) calls this approach "creative evaluation".

consequences of a technical change suggested by an evaluation, when the whole is completely altered by the result? For as our technical capabilities expand the challenges to our conceptions of ethics do as well. Technical change requires adjustments in values and emotions, decisions between "right" and "wrong". The technical ability to create many alternative forms of the same data all of a sudden create value choices, choices that are only resolved through deliberations on "fairness" or "rightness". It is questionable whether people will embrace change if they are uncomfortable with the ethical challenges they will have to face, or if they are uncomfortable with their own ability to deal with such challenges fairly. The holistic conception of organizations also suggests that to employ the IPS in a systematic, cyclical way (i.e., once every three years) just to complete the organizational learning loop is not true formative evaluation. For evaluation must be perceived in the holistic organization as the fundamental skill of organizational renewal; it must be seen as the ongoing renovation skill of the holistic organization. Just as a physical every three years without some form of interim self-evaluation (i.e., symptom awareness, self-examination, etc.) will not keep one healthy, neither will a singular evaluation event keep the organization healthy. This aspect of the IPS did not receive enough emphasis, and was one of its major limitations in practice.

The holistic nature of the Information Profile System also contradicts traditional bureaucratic structures. It is interesting to note that the evaluation project was probably most effective in the district that had an organizational structure and administrative philosophy most consistent with this holistic concept, even embodied in the district slogan: "Together We Learn". It appeared to be least successful in the one that tended toward a hierarchical, role-specific organizational model. Is a possible limitation to the failure of past formative evaluations the simple fact that if organizational learning is required for organizational growth, the bureaucratic structure forces formative evaluation into a summative control mode that cannot be overcome without organizational reconstruction?

The final limitations to evaluation effectiveness were noted earlier. These are a function of resources (personnel, money and time); timing; individual understanding, skill and expertise; and circumstance. To some degree they are a function of some of the other qualities mentioned earlier: vision, will, and

ownership; because in their allocation they become a manifestation of value priorities. As representations of value priorities, they give some credence to the leader-substitute construct, for they become aspects of structure which once in place, inevitably influence future actions. Further discussion on this concept will follow in the next section.

One is left, at the conclusion of the research, with a number of questions, additional to those enumerated above, about the role of formative evaluation in districts. Does a real commitment to this vision exist? What is the relationship between vision and will? Does vision precede a claim to will, in that the former is required to achieve the latter, or does will, in fact, help to create vision? What role does understanding, knowledge, and expertise have to do with "will", and with the development of vision? And, finally, what is the nature of the relationship between evaluation data and the decision making process, both conceptually and practically, when one is construing this data at an organizational level?

Evaluation in an organizational setting is like judgment in real life: it is the highest level of critical thinking. It offers great promise for organizational growth, but also requires the highest intellectual and ethical skills. It may require a complete reconstruction of our image of organizations and the meaning of organizational life. Absence of these abilities and of will inexorably dictates severe limitations on the potential success of evaluation.

Leader-Substitute Theory

This study looks at the role of evaluation in district administration. Insights into the nature of administration have been provided through observations of the dynamics of administration in action, i.e., leadership, in two instances: (1) the evaluation project itself; and (2) before and after the evaluation committee's work in each school district. What has been learned from these two situations?

First, it appears that the evaluation project itself depended upon its vision being understood, on effective management of meaning skills, on skill in group decision making, on the creation of teamwork and group commitment, or "will", on ownership, and on an understanding of how ethical issues arise from technical ones. Second, it also appears that specific evaluation methodologies (i.e., stakeholder collaborative process, goal-free inductive

development of key indicators, comparison of indicator results to the official values framework; "creative" report writing) all can have some positive impact on these qualities district-wide. It appears that the methodological skills of administration or leadership are thus fundamentally related to the process of evaluation, and district evaluation here had some potential for enhancing those skills.

How do these administrative/evaluative dynamics provide insight into the leader-substitute construct of leadership? First, it must be noted that the leader-substitute construct is an outgrowth of the situational leadership theory. There is evidence in the study to show that the existing dynamics within the district--the situation--influenced the evaluation's success greatly. These situational elements appeared to be (1) the purposes and needs envisaged for the evaluation; (2) the personal leadership skill of the supervisor and stakeholders involved in the project; (3) the existing leadership culture, i.e., customs, expectations, and attitudes of district personnel to decision making; (4) existing policies, processes and procedures, such as existing mission statements, or planning policies; and (5) situational issues such as negotiations, time of the year, resources available, etc.

Another precept of the leader-substitute theory is the belief that the power to influence, and therefore to lead, is a function of the interaction between the formal leader's personal ability to influence and the situational elements he is forced to deal with. The leader-substitute construct simply extends this conception to argue that situational elements in themselves are able to influence and lead; that the individual's influence is therefore enhanced or neutralized by contextual leadership factors. Kerr and Jermier's work (1978) and the follow-up studies by others (Ford, [1981]; Gamoran & Dreeben, [1986]; Freeston, [1987]) indicate that there are four main sources of potential influence: the formal leader's own skill, the nature of the task being attempted, the characteristics of subordinates, and existing elements of structure in the organization. The extent to which these are consistent with each other and the overall purpose will determine to a great extent the effectiveness of leadership initiatives to achieve it.

The leader-substitute concept was employed as a guide to the development of the IPS. The idea was to guide district leaders in the creation of a formative evaluation process in their district; which would be a leader-substitute itself.

For this reason the researcher concentrated on minimizing personal influence, and maximizing the influence of task, subordinates, and structure in the development of the process. Evaluation was made the overarching task of all committee meetings. Structures were employed to ensure rational evaluation (i.e., judgment applied to criteria of an ideal) extending from the values framework itself to criteria both for judging a key indicator and evaluating data quality and significance. Subordinate characteristics were employed in that the leaders of various subgroups were invited to become participants in the stakeholder groups, and decisions made by that group became a function of their group learnings and personal intuitive judgments. The researcher then took on the role of facilitator, and tried to let the tasks, structures and subordinates influence the evaluation's completion.

The question to be asked at this point then is whether the personal attributes of the researcher and other administrators were the principal factor determining the success of the evaluation, or did other factors have a supporting or inhibiting role. Did the task(s) of evaluation, the existing structures within the district (i.e., existing policies and procedures; formal role definitions) and the imposed structures (i.e., criteria upon which judgements were to be made, rules re brainstorming, decision making, timelines), within the duration of the project itself, and the characteristics of subordinates (abilities, culture customs, traditions, or beliefs due to secondary and informal role commitments) also influence people either in the direction of the ultimate purpose, or away from it?

It should be clear from the description both of the interactions found within each case study, and of the varying degrees of success of the project dependent on issues such as timelines, report limitations, subordinate skill, etc. that aspects of the task, of structure, and of subordinate characteristics did have major influences on project success. Colin Scott articulated the concept well when he stated:

...I felt myself more and more willing to let the process guide...I didn't feel subordinate to any individual, I felt subordinate to the process...I felt the process controlled me....the process was very useful in that it didn't center on anybody,...but...was very good in subordinating a strong person like myself.

Or, as Jason Stone stated:

When you are sharing in a group...there is a certain amount of control that everyone places on it and by setting the initial parameters you are somewhat controlling it...for example, the structure using our mission statement...could be considered to be controlling...

Certainly individuals in the Inlet and South Coast evaluations felt controlled by the Ministry questionnaire which was a structural component of the evaluation. Also, when stakeholders in the Interior evaluation were asked to identify the factors that were most influential in "the success and integrity of the evaluation," the three with equal scores at the top of the list were: (1) Q-sort that allowed equal input in decision making; (2) senior administrators treated committee members with respect and consideration; (3) the use of judgement and evaluation at each stage of decision making throughout the process. The first can be considered an influence of structure; the second a personal influence; and the third an influence of task. It is important to note as well that the factor with the lowest rating of the nine presented (see Appendix G3 for the total list) was "senior administrators knowing what they wanted and persuading people to do as they wanted." In addition, teacher commitment to the BCTF and its unionized goals appeared to be a subordinate characteristic leading teachers away from active cooperation with follow-up procedures in each district.

The leader-substitute theory suggests that administration exercises its leadership responsibility through both the administrators and the leader-substitutes that exist in the organization (see Figure 8.3). One part of the administrative unit is constantly in operation, but shifting in locus (see Figure 8.4).

Administrative Effectiveness

Hodgkinson (1983) makes the point about administration that it "can be done well, done indifferently, but it cannot not be done at all" (p. 196). One can conceive of poor administration, ineffective administration, and good administration; however, the use of the same terms to define leadership are not as easily accepted. Coombs (1988) for example, makes the point that leadership, as a status position, can be evaluated in this way; but leadership, as

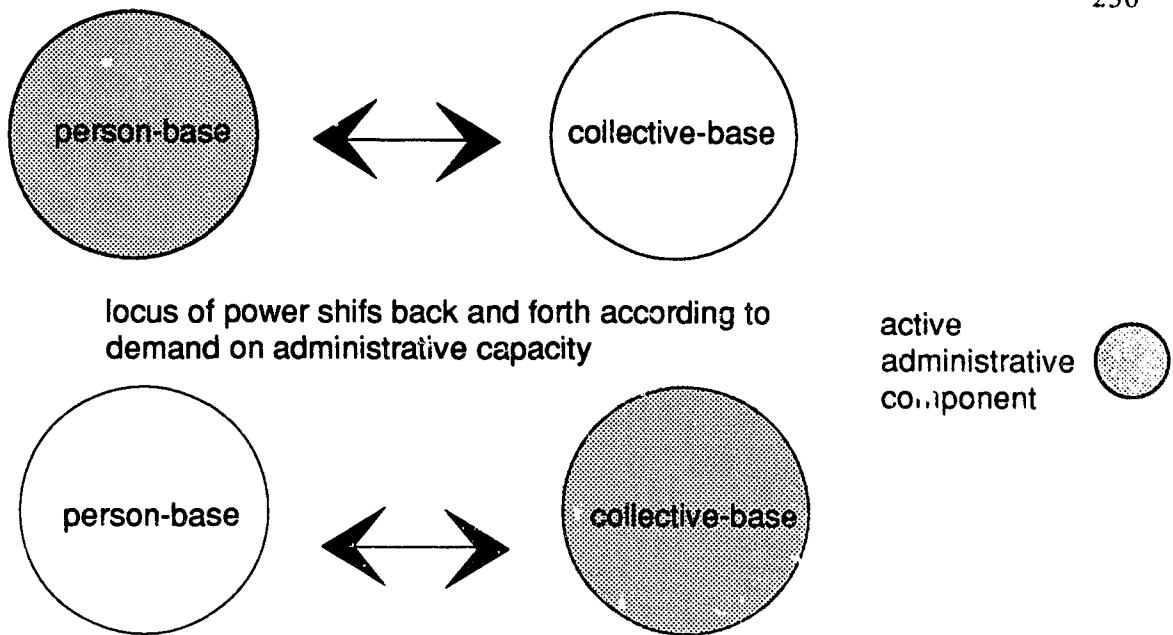


Figure 8.3: The components of administration: person-based leadership plus leader-substitutes.

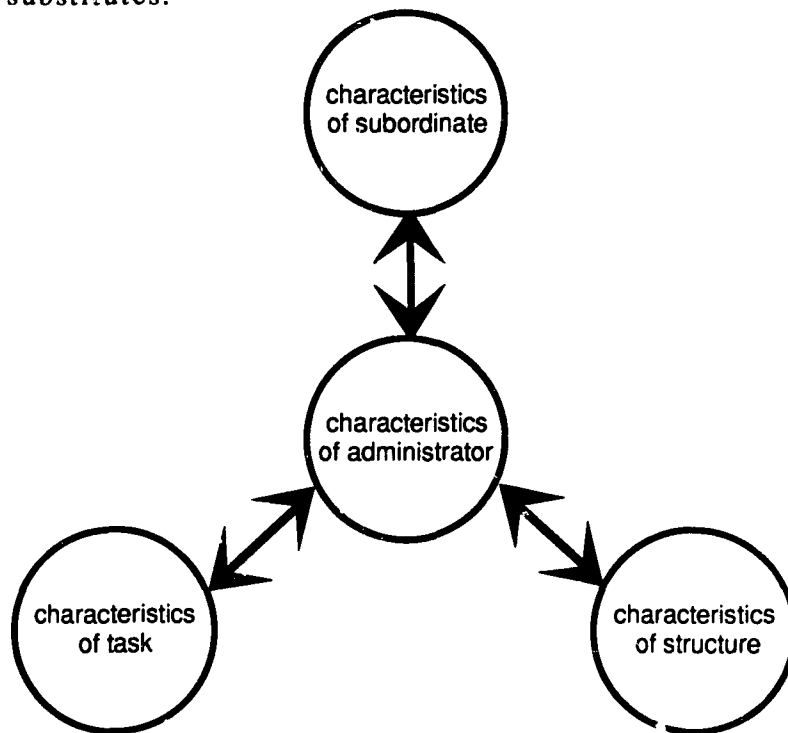


Figure 8.4: The shifting locus of administrative action: constantly in operation.

the capacity giving rise to the evidence of leadership, carries with it a strong normative sense of approval. Defining administration as the composite of leader plus leader-substitute activity explains how administration can be seen as a negative factor, while leadership itself, need not be.

The argument can be stated as follows. The perceived "ebbs and flows" of "status" leadership is simply due to the existence of leader-substitutes. Existing qualities of subordinates, task expectation, and structure in the organization carry on, as Foucault stated, the continually operative leadership that is required to focus people and their behaviors on a day-to-day basis. In these circumstances, the position leader has four options open: (1) expand leadership energy in a manner designed to enhance existing structures or operations; (2) expend leadership energy in a manner designed to reduce the effectiveness of existing structures or operations, and create new ones in their place; (3) expend no energy at improving or reducing the effectiveness of existing structures or operations, but rather utilize leadership skills to deal with problems or issues that have no structures or processes currently in place to deal with them; or (4) expend no energy at all: let the system operate without the exercise of person-power from the formal leader. Regardless of the leader's decision, administration is taking place; yet only in the first three instances is leadership occurring, i.e., person based leadership. However, whether administration is positive or negative in its manifestation is a function of how the person leader's efforts interacts with the leader-substitutes, and the impact that has on the overall strength of the organization.

Figure 8.5 illustrates this. The implication is that when administrator behavior complements the leadership potential of the organization's leader-substitutes, the leadership capacity of the organization grows (i.e., quadrant 4). This can happen in two ways. It can be through the exercise of the formal leader's personality and charisma, whereby existing leader-substitutes are infused with greater energy and effectiveness; or, it can be when these qualities are used to fulfill needs in areas whereby leader-substitutes do not yet exist. In the latter case, the administrator may exercise organizational skill, building leader-substitutes where none existed in the past, or simply deal with the challenge in a manner that "models" effective leader behavior. On the other hand, a person-leader's penchant to act when perfectly sufficient

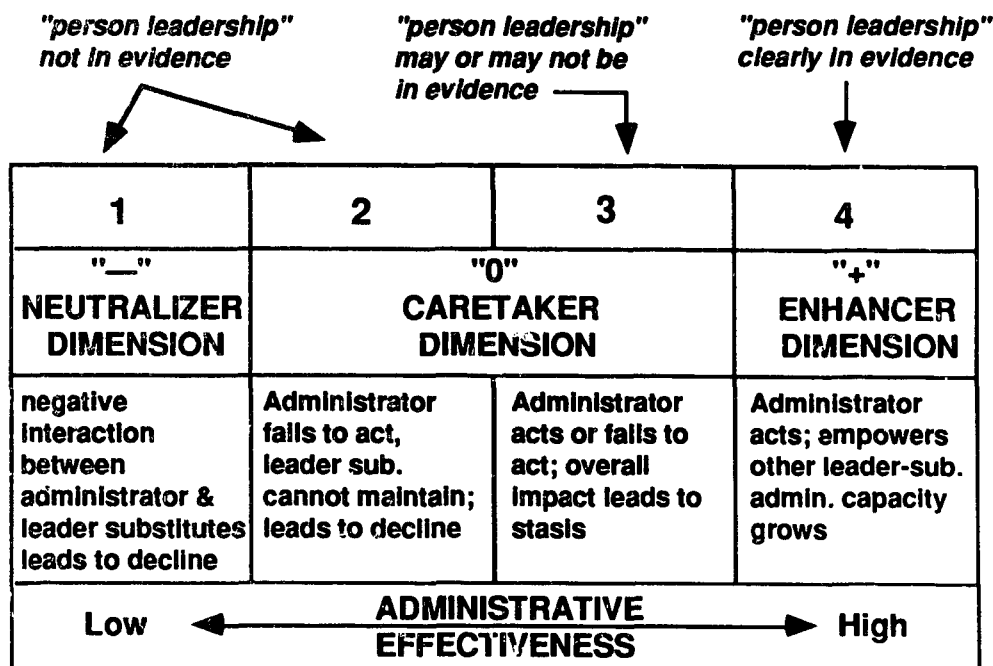


Figure 8.5: The dimensions of administrative effectiveness.

leader-substitutes already exist can cause organizational decline. It may mean energy is not placed in areas of more fundamental need, or if done in an insensitive or irresponsible manner may lead to the disempowerment of those existing substitutes. In this circumstance, the administrative capacity of the organization is weakened (see quadrant 1). Similarly, the failure to act can have two effects: (1) stasis--a state whereby the organization is neither strengthened or weakened; or (2) decline--whereby the organization is weakened because the infusion of energy needed by the formal leader is not forthcoming (see quadrants 3 and 4).

In conclusion to this section it should be stated clearly that the leader-substitute construct is a concept of leadership that is more reflective of the subjectivist paradigm of inquiry than the scientific one. As Clarke (1985) points out situational leadership theory exists in recognition of the confines of context, and the necessity of gaining understanding through studies of interactions and relationships between components rather than through understanding of each separate piece. It implies that leadership influence is a dynamic interaction between indeterminate variables. Clarke (1985) also points out that the emerging paradigm also has given rise to two other related

shifts in understanding: first, a belief in the concept of "loosely coupled organizations"; and second, in acceptance of new mechanisms for effective leadership in these nonbureaucratic types of organization. It must also be stressed, however, that this is not an empirical endorsement of the theoretical validity of the leader-substitute theory of leadership. In many ways it simply indicates that the leader-substitute construct is useful in understanding the dynamics of administration, and it does so in the following ways.

First, one should not attribute the success of any enterprise solely to the ability of the formal leader. Personality or style must complement the leadership needs of the organization, but the impact they have will be dependent on how personal style meshes with other leadership influences that are existing in current structure, task assignment, and cultural characteristics and personal abilities of subordinates. Much of the recent literature in administration emphasizes the importance of the latter, suggesting that the cultural values of subordinates are a major factor influencing organizational success. However, the leader-substitute construct suggests that not only should subordinate characteristics be looked at as a source of value commitment. The existing policies, procedures, and task responsibilities of people within the organization should also be examined. They represent value priorities of the past; some of which are still of high priority, and others little more than artifacts impeding progress.

That there is further validity to the leader-substitute construct is shown by the fact that value commitments inherent in structure and task are often unquestioned, or "operative in ignorance." If such influences do not reflect the value priorities of today, the administrator who does not recognize the influence they are having may have new initiatives founder without understanding as to why. It also points out the integral importance of evaluation as an administrative function. For it is through evaluation that one discovers existing value priorities. It is also through evaluation that one can examine the ethical congruency of values inherent in existing structures and tasks with those values envisaged for the future.

Another merit in the leader-substitute concept is its suggestion as to how one can minimize reliance on personal authority by building influence into aspects of task, structure, or subordinate characteristics. The latter are reflected in the concepts of ownership or empowerment, whereby subordinate

values, skills and abilities are bound to the organization's goals. Task and structure suggest how effective meeting agendas, timelines, plans, and decision making strategies developed in advance can assist an individual in moving a group toward or away from a predetermined goal. Plans employing these stratagems can become a mechanism for the engendering of will and the commitment of people to an organizational purpose. For example, what value does a superintendent's agenda for a principal's meeting suggest when all the tasks are trivial paper management? What value does a decision making process done in bureaucratic isolation reflect as opposed to a decision making process done in the context of a stakeholder collaborative group? If these tasks and structures are designed in an ethical manner, then they can reduce the administrator's need to employ personal persuasion and skill to come to appropriate conclusions. For policies, tasks, processes and procedures are little more than values parading as formal statements of intent; the more they are congruent with the values of the organization the more likely that the organization will operate in a purposeful, rational fashion.

The leader-substitute concept does suggest why management of meaning is so difficult in education. Educational leaders have often been immersed in the educational system since they were six years old. How aware can one be of the values embedded in structures and tasks that one has experienced from such an early age? It is like fish being unaware of the water in which they swim. Parents and others suffer from the same problem. How are their present values affected when they walk into a school and are reminded by the teacher and the classroom structure of values of the past? We have all heard so much about "back to the basics"--how impervious is this "cultural belief" to change?

These observations help explain why the issues of vision, management of meaning, ownership, and will were psychological factors in the three case studies. Each of the above is a values resolution and congruency-seeking process. Each, as much as each stage of the evaluation project, and the purposes of the evaluation project itself, is an attempt to sort through and identify values of different priority and significance, both to individuals, to society, to sub-organizations within the organizational whole and the organization itself. These values are not only explicit in the statements of people and their actions, but also implicit in the situational context of the district. Ultimately, administration, or leadership (Hodgkinson, 1978), is little

more than a values resolution process; and formal evaluation at the district level simply an attempt to impose some degree of order and control on what would otherwise be a intuitive and intensely personal activity, rather than a formalized institutional learning process.

We have now come almost full circle: issues of value resolution and the emphasis they place on evaluation, vision, management of meaning and ethics point out once again the centrality of philosophy to administration. This once again directs us to the writings of Hodgkinson (1978, 1983); Barnard (1938) and Greenfield, (1975, 1979, 1983). Hodgkinson's discussion of differentiation of values of different potency and moral strength in his value paradigm has potential as a guide to value resolution in evaluation; it is shown in Appendix G3.

Barnard's (1938) characterization of the ultimate administrative skill as consisting of two abilities: (1) creating ethical solutions as change forces new contexts on old values; and (2) the importance of will in administrative effectiveness is most apropos to the findings of this study. He suggests that time, perspective (i.e., level within the organization), and subordinate characteristics all contribute to these solutions. The research conducted in this study suggests that evaluation of the organization as a holistic entity, especially when grounded in cooperative evaluation strategies may help clarify ethical issues and also assist (through vision-building and teamwork) in creating higher levels of organizational commitment. It also suggests that success in administration depends to a great extent on the ethics and will possessed by the administrator. The latter quality may be stimulated by utilizing aspects of influence other than personal charisma; and existing cultures, structures and tasks may detract from it. The leader-substitute construct also supports the emphasis a number of writers are placing on organizational culture.

The study also lends credibility to Greenfield's (1975) prolegomena for the study of administrative behavior in organizations. Of particular import are his statements that "organizations are expressions of will, intention and value"; "that facts do not exist except as they are called into existence by human action and interest" (witness the issues related to data analysis and representation); "organizations are arbitrary definitions of reality woven in symbols and expressed in language"; and "That organizations expressed as contexts for

human action can be resolved into meaning, moral order, and power." Each of these statements accurately described the issues that came into play as determinants of the success of the IPS in each district case. A further statement--"man acts and will judge action"--highlights the importance of utilizing organizational evaluation skills to inform administrative action; not for administrators alone, but for all of its members.

Further Implications

The discussion on the limitations of evaluation use and leader-substitutes provides a logical explanation both as to why formative evaluation, in the form of the IPS, was suggested as a mechanism for organizational improvement--as well as why it was limited in its success. First, the project was generated when senior administrators, at the district and Ministry level--felt that control of education was slipping away; i.e., their careers were threatened, public perception of education performance was low, and controversy surrounding education was affecting educator's ability to work. This is an illustration of the meta-value principle described by Hodgkinson (1978); in that two fundamental values asserted themselves: a desire to regain control; and a desire to improve. A formative district evaluation model was seen to be the key to both. But formative evaluation--particularly the holistic model suggested in this study--priorizes improvement before control. And, in pursuing improvement it employs two mechanisms that, in the short term at least, appear to minimize the administrator's personal control: (1) it increases the potential or possibility of value congruency through improved knowledge; it binds the administrator to negotiated actions that tie his hands and diminish his personal ability to influence; and (2) it builds influence into structures, tasks, and subordinate characteristics that create a system that an administrator cannot completely control through personal qualities alone. However, these two factors diminish personal administrative control, but expand overall control. If an administrator can commit to improvement before personal control, the evaluation model will be effective; if, however, personal control is truly the motivating factor, formative evaluation will be allowed to lapse and old control mechanisms will reassert themselves.

Formative evaluation's power can also be its major weakness. It may be

postulated that educational organizations might sometimes be said to be slowly evolving away from the bureaucratic model, into what is termed the "cultural image" model: "The cultural image uses processes and symbols, has a broad and internally forged definition of purpose, and utilizes a 'child-as-client' orientation" (Teddle, Kirby and Stringfield, 1989, p. 234). But this change does not come easily, and people do not necessarily adapt comfortably. The currency of control in a bureaucratic organization is authority, associated with personal status, specialized role responsibility, expertise, and control over the granting of rewards and punishments. In a sense the formal leader becomes the embodiment of organizational purpose and as such his personal authority must remain paramount.

Formative organizational evaluation would suggest very much the opposite. The embodiment of organizational purpose becomes a set of ideals or values. Control is less a function of authority and more a function of value congruency. Value congruency is sought through diminishment of hierarchical status, creation of trust and teamwork, and placing influence in tasks and structures that embody democratic, participatory principles. The currency of influence is information, not role or status. The suggested purposes of evaluation--improved accountability, maintenance and gain of public and professional support for education, and improved decision making --are little more than results attributable to an "educated" organization. Organizational improvement through structural education replaces formal control by the leader as the source of organizational change. If a formal leader can see the vision, accept loss of personal control, discover the leader-substitutes that reflect adherence to the value of bureaucratic control and find mechanisms to overcome them, and be possessed of the will and ethical skill to deal with challenges the change will present, then he or she can restructure the organization into a true learning entity in which formative evaluation will be embraced; if not, formative evaluation will have failed. Thus, the potentialities of formative evaluation are both its main strength, and its main limitation.

The above observations derived from this research methodology prompts one to suggest that the role of educational administrator is more analogous to that of a theatrical director than factory manager. A director reacts to a script, and intuits a theme, or vision, that will guide the production. In

converting that vision into a production, he employs the skills of ownership and evaluation. All endeavors are directed at improving the production; personal control is only employed where necessary to accomplish the meta-value of improvement. Ownership is also important in that each actor must somehow mesh his role and his personal values with that of the production's values; and also in that the audience--the client--must also value the final result. To achieve this he alters his vision and creates, with actors, a shared one. Evaluation is the tool he uses to do this. Formative evaluation is employed in progressing step by step towards the final product. It is systematic; for each rehearsal, each performance, is judged according to its ability to reach the ultimate vision. It is employed in a forward seeking, improvement oriented way.

The director of a play also acknowledges the power of leader-substitutes. He knows that the aspirations, needs, and skills of each actor influence the ultimate performance; he also knows that the structures such as walk, costume, and props can create within the actor or actress an understanding of the character as envisaged by the director yet in keeping with the play's theme. Blocking¹¹ (another structural concept) and the set contribute immensely to the production's ultimate success, but are specific to the particular production--they manifest the values unique to that production, and will be dispensed with upon its completion. Evaluation becomes the fundamental directorial skill; it is the source of judgment whereby each prop, each structure, each task is judged as complementary to the production or not; and whereby individual needs are connected to those of the production's ultimate vision. And finally, the major determinant of the director's success is whether the play had the effect on the audience that they wished it to have; similarly, in education, administrative effectiveness should be determined by student, parent, and public perception as to the effect it had on student growth.

¹¹ Blocking is the process by which the director places actors and actresses within the stage set so as to gain the maximum visual impact for each scene; it is usually done consistent with generally well-accepted blocking principles.

Summary

The three case studies indicate that evaluation--and administration--are little more than values resolution and consolidation processes. They also imply that effective administration and leadership in a loosely-coupled organization depends on evaluation, and evaluation success is predicated on administrative commitment to improvement rather than career success. The research also suggests that a leader's ability to design processes and procedures, that direct organizational members into problem solving, and then to enlist commitment to the pursuit of solutions is the key to evaluation success. Furthermore it is possible that a bureaucratic or rationalistic approach to administration militates against this possibility, because it does not allow for truly meaningful sharing of value dispositions and knowledge. A holistic evaluation process may be a positive step towards restructuring of traditional administration; but as an event in a systematic, ongoing bureaucratic process, it is likely to have little substantive influence in the short term. The study also suggests that alternative metaphors for administration may be appropriate in viewing the skills and abilities required in the administration of education.

One conclusion is hard to escape, however. Clearly the IPS will have very limited effectiveness in the short term on a district's operation. The desire to employ it, however, may represent an awareness on behalf of administrators that a true change in the nature of leadership is possible; a change that might slowly enhance the role of evaluation as a key component in effective administration. Perhaps it is wise in terms of the timeline for this change to be guided by the thoughts of the parent representative on the Inlet stakeholder committee, thoughts that seem duly appropriate on this occasion:

I see the evaluation as a process of awareness...if I become aware, there will be a ripple effect. We have to reach out and make things happen. If we can transmit this to others, then it will grow. This is how change is made in the system; it is a slow process.

CHAPTER IX

Summary, Conclusions, and Recommendations

The first section of the chapter summarizes the study, including the background, the problem, the design of the study, the methodology and the findings. The second section of the chapter details the implications of the findings.

Background

The present study came about because the researcher wanted to know how formative evaluation could be constructed to assist district officials with educational administration, what limitations might exist as to its ultimate effectiveness, and what insights the employment of the evaluation process provide into a better understanding of the nature of educational administration. Steps had been taken to develop such a model in British Columbia, called the IPS. Research of the literature indicated that in order to answer these questions it would be necessary to examine closely the literature dealing with the generation of the IPS itself, administration in loosely coupled organizations, values in administration, evaluation, and holistic evaluation.

From the literature pertaining to the generation of the IPS a number of basic purposes and procedures and principles were identified as relevant to the evaluation model. These were:

1. Purposes:

- a. improvement of planning and educational performance.
- b. motivation and gain of public and professional support for education.
- c. accountability enhancement.

2. Procedures:

- a. the use of a values framework, in the form of the Four-Goal and Six-attributes (*q. v.*), to guide the process.
- b. the use of key indicators of performance to measure achievement of educational success.

- c. the use of questionnaires of stakeholders to determine perceptions of educational success.

3. Principles:

- a. flexibility to adjust to each district's individual context.
- b. a cooperatively developed, collected and accepted body of data.

Research associated with a holistic world-view suggested viewing the educational organization at the school district level as a loosely coupled organization. In turn this conceptualization suggested examining the organization as a holistic entity rather than a mechanistic one; and emphasizing the role of values in organizational decision making. In this context administration was seen to have the function of facilitating organizational problem solving. Various theories of administration were examined, and the leader-substitute concept of administration--whereby the operational influences were found not primarily in the personality and style of the formal leader, but within characteristics of task, structure, and subordinates found within the school district--seemed to be consistent with the holistic vision of organizations earlier presented. To succeed in problem-solving the literature suggested that the various facets of administration would be successful if they worked together to create a vision for the organization, to develop ethical relationships through participatory decision making, to manage meaning in the organization, and empower organizational members. Formative evaluation was seen to be the key to development of these administrative skills.

Examination of the evaluation literature suggested that the methodology to answer academic and research questions should rely on qualitative rather than quantitative models of evaluation. Principles relating to the naturalistic form of evaluation were chosen to guide not only the district evaluation model but the study itself. The concepts of multiperspectives, stakeholder collaboration, utilization-focus, creative evaluation, goal-free inductive analysis, and reliance on the intuitive evaluation skills of the human agent were then meshed with the utilization focus of the holistic model of large system evaluation. A holistic model, employing the givens of the IPS--i.e., questionnaires, the values framework, key indicators and evaluation reports--

was then structured according to the principles of evaluation mentioned earlier.

In essence, the literature suggested that qualitative, philosophical administrative principles, combined with qualitative evaluation strategies would result in a model of formative evaluation that would most likely achieve the goals established for the IPS. The literature suggested this because both approaches focus on values resolution and clarification; and a congruent value focus amongst all the sub-groups of a loosely-coupled organization enhances effective decision making.

Study Purposes

A major assumption inherent in the development of the IPS is that a formative evaluation model--properly constructed--should be able to accomplish the goals envisaged for it. Yet the literature also suggests that there is little evidence that evaluation has achieved in practice what theory promises. This problem generated three purposes for the study. A first purpose became to design a holistic, formative evaluation model that would enhance administrative capacity at the school district level--a model that administrators would be prepared to use because of its practical and theoretical value. A secondary purpose was to design a study that would lead to further understanding of the factors that delimit evaluation's effectiveness when in the hands of administrators. A third purpose was to use the insights and experiences of the study to enhance our understanding and insight into the dynamics of both evaluation and administration, specifically the leader-substitute construct.

Design and Methodology

In keeping with the study's purposes, action research was employed in three school districts undergoing case studies in the use of the IPS. Naturalistic evaluation methods were used to evaluate and meta-evaluate the effectiveness of the model, and to gain insights into the limitations to evaluation effectiveness, to the dynamic relationship between evaluation and administration, and to the leader-substitute construct itself.

Pursuit of the study's first purpose--development of a formative evaluation model for school districts that works--was done in two stages. First, the

researcher carried out action research while employed as facilitator of the evaluation project in each pilot district. Attempts to improve the evaluation's procedures and processes were guided by the given aspects of the IPS, by questions that were posed at the end of the first pilot test, by strategies the literature suggested might provide answers to the questions, by situationally determined administrative responses, and by the limits imposed by the districts themselves on the project. The second stage was to critique each case's effectiveness, and to meta-evaluate all three cases' effectiveness. The meta-evaluation was done in two steps: first, by answering the four questions posed at the beginning of the study; and second, by examining the varying degrees of success of the cases in terms of achievement of their administrative objectives. As the cases were completed in succession, some lessons learned in the first two were applied to the last. Secondary purposes, i.e., looking for insights into limitations on evaluation effectiveness, and the dynamic relationship between evaluation and administration and the leader-substitute construct, were achieved through the naturalistic inquiry conducted in the three cases.

The main source of data for the study came from interviews of stakeholder team members who participated in the cases. Supporting data came from meeting minutes, the researcher's logs, notes taken by stakeholder's themselves, notes provided by Ministry employees who assisted with the projects, newspaper stories, letters, and district evaluation reports. Data were content analysed and color-coded according to five main areas of interest: (1) background and context information; (2) description and/or evaluation of the project's methodology; (3) evaluation applications; (4) perceived limitations to evaluation effectiveness; and (5) insights into the leader-substitute construct. Findings with regards to (1), (2) and (3) were presented as individual district case studies, organized around a single conceptual framework that revealed itself through analysis of the data. Meta-evaluation of the model's effectiveness was presented as a cross-case study analysis, building upon patterns found across the three studies. Similar, patterns between districts were sought out to provide the insights required to make sense out of the data categorized in (4) and (5). These understandings were presented in Chapter IX.

Findings

First, as an "ideal" values framework for the evaluation, the Ministry's 4-Goal and 6-Attribute framework, with minor modifications, met with approval in all three districts. It might serve either as a guide upon which a district's mission statement could be constructed, or as a statement of the district's values.

Second, key indicators of performance--determined locally, measured locally, interpreted locally, and represented locally--were used to measure each district's educational performance in all goal and attribute values. The key indicators were most powerful in management of meaning when they represented a goal or attribute in three different forms: first, as a quantitative measure of performance (i.e., a statistic representing a potential range of achievement); second, a qualitative measure, i.e., a perception check of public opinion as to educational achievement; and third, a demonstration measure--an example--of the performance criteria indicated by the other two measures (a working definition of the qualities of a good key indicator is found on page 206). Key indicators are inductively designed and temporally meaningful; they will become more powerful when measured over time.

Third, techniques such as brainstorming, the Q-sort, and the stakeholder collaborative process appeared to create a sense of ownership for the evaluation project, as does control over data collection, interpretation, and report writing. They also enhance the potentiality of reaching negotiated judgments as one moves from the specific to the general in evaluating district performance.

A fourth finding was that surveys appeared more meaningful and acceptable when designed and conducted (1) with significant stakeholder input; and (2) by professional pollsters. This finding, combined with the previous ones, suggests that standardization of either key indicators or survey questions from the Ministry is undesirable; any such consistency should develop in an inductive manner from below. However, due to the need to employ comparison for understanding, a process to accomplish this is also desirable.

Fifth, it became apparent that careful attention to timelines and timing were important for the evaluation's success. The project that was attempted during the school year was much more successful in data gathering than the two summer-time projects.

A sixth issue relates to the state of data collection procedures within each district. The system requires data about student and organization performance. Yet no such data pool exists. A startling paucity of requisite information, at least data that is retrievable by others, characterizes schools and school districts alike. Related to this issue is the unease educators have about interpretations of this data, and its true meaning.

A seventh finding relates to the method of reporting evaluation information. Audience specific reports are desirable. Information presentation through appeal to various learning styles may be desirable. The effect of video reports, for example, on the use of information should be examined further. Report styles which match initial purpose should also be examined.

A eighth finding relates to data management and meaning. Finding "reality" or "truth" in educational information provided by the three districts was almost impossible. The information provided was more, as Greenfield would say, the result of "human intention and action", consensually--but subjectively--determined. This created problems with maintaining a clarity of purpose on pursuing the goal of district improvement, and with ethics of information management for administrators.

Finally, much greater efforts must be placed on groundwork and follow-up if the IPS is to be used effectively. Indeed, failure to create a "bridge" from evaluation to decision making through the use of follow-up procedures designed in advance was a major weakness in the IPS.

Analysis of the three case studies suggest that the IPS methodology was ethical, empowering, and informational for the stakeholder teams themselves. Transfer of these attributes into improved district accountability, planning, and gain of public and professional support for educational improvement varied greatly from district to district. Perceived success was partly due to the evolving sophistication of the evaluation model itself, and partly due to the fact that the System ignored the most important component of effective evaluation: follow-up.

As to accountability, the IPS heightened awareness and dialogue amongst participants, but only when organized follow-up processes were followed did the apparent success extend to other district members. Use of evaluation information to enhance decision making through development of a long term

planning strategy was essentially a failure. No districts designed a report to present information for planning, but the evaluation information was incorporated into existing planning structures. Although all districts indicated a desire to construct follow-up evaluations and use of key indicators for district monitoring, no commitments of an unequivocal nature were made by any of the district management teams.

A pattern of results similar to that with respect to accountability and planning was discovered when gauging whether the IPS created an increased sense of teamwork and commitment amongst staff and public in the district. Evidence existed to support an affirmative answer for members of the stakeholder teams, but not so for the remainder of the district's employees. This is likely due to the fact that the case study was an event, not an institutionalized process, and that the anticipated goal is long term, rather than a short term.

In summation, it appeared that the reasons for the limited effectiveness of the IPS were (1) the political importance the project assumed within the district; (2) existing patterns of leadership and style within the district; (3) the characterological ability or style of each actor in the process; (4) the interaction between these qualities and the methodology required by the evaluation project; (5) the follow-up processes employed in the district; and (6) the resources and situational elements found in the context of each district.

General Limitations

Limitations to evaluation effectiveness in this study were basically of two kinds: (1) evaluative, i.e., as the design and method meshed with the political and organizational realities of each district; and (2) administrative, i.e., Having to do with philosophical and personal expertise demands placed upon educational administrators.

Evaluative Limitations

Evaluative limitations were of two types: traditional limitations, and what can best be termed as "other limitations". In the former category rests apprehension surrounding evaluation due to its potential use as a control mechanism, or for political maneuvering by one political sub-group against

another. In the "other" category were three main subsets: (1) psychological limitations; (2) holistic limitations; and (3) change limitations.

Psychological limitations confirmed that the greater the trust that exists amongst organizational members, the greater the likelihood of a meaningful evaluation. Second, the more evaluators were motivated to improve, the greater the likelihood of success. Holistic limitations refer to the complexity of holistic evaluation itself (i.e., getting consensual value judgments), and the incompatibility of holistic evaluation with traditional bureaucratic role responsibility. Change limitations refer a natural resistance to evaluation because it is a harbinger of change. A strong possibility exists that this is due to the difficulty of envisaging the impact of change in a holistic, indeterminate organization. Related to this is the problem of determining what is good information versus bad information, and how best to present this information so as to relate in an evaluation report, the "real" but at the same time, intensely personal essence of the school district's strengths and weaknesses?

Administrative limitations are both philosophical and practical. Philosophical limitations refer to the issues of vision, will, and ethics in management of meaning; all determine the extent to which various educational leaders are able to perceive the true nature of formative evaluation. Indeed, holistic, formative evaluation suggests a vision of a new "learning" organization, one in which the administrative persona is diminished, while administrative character is enhanced; and in which the bureaucratic infrastructure is transcended with respect to ethics. Practical administrative limitations refer to the ability of district administrators to design follow-up procedures to link the evaluation with its purposes, as well as the ability to employ tactics, strategies and skills to create value congruencies in the organization.

Leader-Substitute Theory

Although the study provided little empirical evidence to validate the leader-substitute construct as the "correct" theory of leadership, it did provide supportive evidence for its relevance within an overall complexified theory of leadership, consistent with the emerging paradigm. It also demonstrated that as a conceptual tool, either for developing an administrative structural

initiative (like the IPS) or for analysing the reasons for administrative effectiveness, it had merit.

Conclusions and Implications for Further Research

In keeping with the constraints inherent in a naturalistic inquiry, conclusions are only generalizable across the three cases analysed in this study. Applicability of the conclusions to other districts or to other contexts depends on the reader's judgment as to the fit between situations.

1. The IPS model exists in a format that enables it to conduct an "event" evaluation of a school district's performance in accordance with Ministry expectations. With modifications it can be used to enhance accountability; without major modifications it will not significantly enhance decision making or maintenance and gain of public and staff support for education.

2. Major modifications necessary for improved performance relate to the creation of groundwork, reporting, and follow-up procedures that are specific to the purpose envisaged for the project. In addition, these modifications should extend the processes to the district as a whole.

3. No modifications or alterations of the process for standardization of key indicators should be considered, unless adherence to that standardization process is voluntary, and the process itself is inductive, stakeholder-based, and goal-free.

4. If the IPS is to be true formative evaluation, then follow-up processes that significantly improve data collection and interpretation, systematic key indicator measurement, regular and effective communication, and regular and consistent stakeholder input must be in place. In addition, and more importantly, the ethics of information management must be defined and clarified for all district personnel.

5. A variety of report formats for the presentation of evaluation information should be experimented with. The impact of multi-media reporting, video reporting, and print reporting should be compared and studied. Similarly, the different mechanisms for information presentation (i.e., graphics, pictorials, etc.) need to be examined in terms of their power and integrity.

6. The major success of the IPS perhaps lies in its educative function. Improved decision making is a secondary accomplishment only possible when

the System is expanded into a true formative mode; the same for motivation and gain of professional and public support for education.

7. Evaluation at the school district level is as much an administrative task as an evaluative task. As such it is limited as much by administrative skill as by its evaluation methodology.

8. A major limitation to the use of evaluation for leadership purposes was the extent to which individuals affected by the evaluation "owned" all stages of the evaluation's progress; the key to overcoming this limitation is to create a climate of trust, ethics, shared commitment and group purpose.

7. A major limitation to the use of evaluation for long-term planning or district gain of support for education was the authentic commitment of individuals to the purpose of improvement. This willingness was in itself, at least in part, a function of previous training in evaluation, understanding of evaluation, individual will, vision--and ability to resolve ethical conundrums, particularly with respect to issues of information management.

8. A further limitation to the effectiveness of the IPS both as a stage in organizational development or as a beginning to the construction of a true formative evaluation process, is the consistency between the basic principles upon which it is constructed and the existing leadership context. It is a holistic model with suggestions of a new conception of organizations and leadership. It includes a new understanding of the nature of change, of authority structures, and of the skills necessary for effective administration. Districts in which these new understandings are operative are more likely to employ it to improve system performance.

The following general statements--not conclusions-- are suggested by the study as worthy of consideration by the reader with regards to the nature of district evaluation, the leader-substitute construct, and administration:

9. Administration is little more than a values resolution/congruency forming task. Creation of value congruency is a function of consistency obtained between subordinate values, values of task, values of structure, and values of the organization.

10. Formative evaluation is an intention to create rational value congruency in an indeterminate and irrational environment. It will be resisted to the degree to which individual psychology and group psychology resists rational change in any aspect of their life; it will be accepted to the

extent to which administrators are able to create an environment of rationality--i.e., an environment of teamwork, ownership, shared vision, and common meaning, committed to organizational improvement.

11. The research strongly suggests the view that school district organizations are as Greenfield (1983) understood them; as creations of human intention, will, and value.

12. The research also supports the point of view that images of organization and administration consistent with the bureaucratic image are damaging to educational organizations (Teddle, Kirby & Springfield, 1989); they are best viewed as learning organisms, devoted to self-actualization (Morgan, 1987).

Questions for Further Research

1. To what extent can a rational formative evaluation process emphasizing organizational improvement overcome irrational fears related to evaluation? A corollary question is to what extent can these fears truly be called "irrational" in a bureaucratic environment? If follow-up to the evaluation is not similarly designed, is there a point to the evaluation?

2. What is the link between irrational emotions and feelings (i.e., fear and trust) and "rational" values? What is the role of both in triggering acceptance of evaluation for change?

3. To what extent does a subject's past experience with evaluation for control delimit their ability to employ evaluation for improvement?

4. To what extent and/or degree can administrative vision and will be altered so as to enhance evaluation effectiveness, and how can this be done (i.e., philosophical training, theatrical experience)?

5. To what extent does the creation of ownership in evaluation create ownership for improvement?

6. To what extent is the commitment to improvement as a result of evaluation a function of inadequate ethics, morals, and values in information control and management?

7. To what extent is an administrator's commitment to formative evaluation delimited by his need for personal authority? and,

8. Does success of the formative evaluation model require commitment to a completely new administrative style and vision of organization?

The study also pointed out the potential significance of the leader-substitute construct to an understanding of the situational theory of leadership. A number of observations and questions regarding this construct are worthy of further study. For example:

9. Does the construct have validity beyond its use as a convenient conceptual tool?

10. How do administrators restructure an organization so as to make structures and tasks flexible enough so as to complement value priorities for the future? Is this done mentally, by simply readjusting one's priorities?

11. Are administrators more effective when they seek out "hidden" leader-substitute influences, and then design a plan for innovation that embeds influence in these characteristics, rather than relying on personal charisma or persuasion?

12. What utility do metaphors such as the educational administrator as artistic director have for understanding the role of leadership in schools?

13. What potential do the values inherent in structures and tasks have to limit or expand the will and vision of administrators?

Further Recommendations

The previous sections have discussed recommendations regarding further study of the major purposes guiding this study. It is possible that the results of this study may have implications for research in other areas as well, particularly as regards to the nature of the educational organization. For example, what impact does it have on organizational structure if administrators choose to view the boundaries of the organization as including members of the public? What impact does this have on management structures within the district? In addition, what impact does it have on organizational structure when one views it metaphorically as a values prioritizing institution, maintaining control through learning and relearning, as opposed to a mechanistic, cause and effect oriented entity?

A second area in which studies related to this one might have influence is with respect to its impact on schools and administration within schools. How does a district formative evaluation model using key indicators of performance relate to school performance? Should district indicators be presented so as to

allow for school comparisons? What autonomy do schools retain for control over decision making if this is the case?

One might also want to consider application of a similar evaluation model to schools. If so, how does one mesh the two? Then too all the questions concerning district evaluation and leader-substitutes could be examined at the school level for further insight.

Conclusion

At the end of the study one is left with the impression that more questions about the efficacy of formative evaluation in district administration have been asked than answered. In terms of the Information Profile System itself, it has received limited endorsement as an evaluation event, with its greatest impact on accountability. As a formative evaluation model much work yet needs to be done so as to create tasks and structures that complement the intuitive daily evaluation decisions made by educational leaders. If such development is done by keeping the limitations to evaluation effectiveness and the leader-substitute construct of leadership as guides to action, and applying them in the contexts of one's own district, then the model has application where one deems it to be of value in organizational improvement.

Reference List

- Adams, G. S. (1964). Measurement and evaluation in education, psychology and guidance. New York: Holt, Rinehart and Winston.
- Alberty, S. & Mihalik, B. (1989). The use of importance-performance analysis as an evaluative technique in adult education. Evaluation Review 13(1), 33-44.
- Alkin, M., Daillak, R. and White, P. (1979). Using evaluations: Does evaluation make a difference? Beverly Hills, CA: Sage.
- Audits of the Government, British Columbia. (1988). Regard for economy and efficiency in the administration of public funds, and for public accountability. pp. 75-120. Victoria: Queen's Printer.
- Ayers, T.B. (1987). Stakeholders as partners in evaluation: A stakeholder-collaborative approach. Evaluation and Program Planning. 10, pp. 263-271.
- Barber, B. (1988). The conquest of politics. Princeton NJ: Princeton University.
- Barnard, C. (1938). The functions of the executive. Cambridge: Harvard University Press.
- Bates, R. (1985) Liberalism, Marxism and the struggle for the state: Prolegomena to the study of public administration. Victoria: Deakin University.
- Bill 19-1987. The labor relations act. B.C. (May, 1987).
- Bill 20-1987. The Teaching Profession Act. B.C. (May 1987).
- Bloom, A. (1987). The closing of the american mind. New York: Simon and Schuster.
- British Columbia Ministry of Education. (1987). Brief to the Royal Commission. (MIC Publication No. ISSN 0823 0714). Victoria, BC. Ministry of Education Printing Office.
- Brown, F. G. (1976). Principles of educational and psychological testing. New York: Rinehart and Winston.
- Bula, F. (1987a, September 10). Teachers, boards hopeful of end to instruction-only campaign. Vancouver Sun, p. 1.
- Bula, F. (1987b, September 10). Anti-union teachers set for first meeting. Vancouver Sun, p. A3.

- Bula, F. (1987c), September 24). Teachers attack plan to release list of names. Vancouver Sun, p. B1.
- Burns, J. M. (1978). Leadership. New York: Harper and Row.
- Calam, J., & Fleming, T. (1988). British Columbia schools and society. B.C. Royal Commission on Education, Commissioned Papers: Volume 1. Victoria: Queen's Printer.
- Cameron, I. (1988). A description of the activities of the program evaluation branch. Unpublished working paper, BC. Ministry of Education, Victoria.
- Cameron, I. (1987). Categorization of districts. Unpublished working paper, BC. Ministry of Education, Victoria.
- Clark, D. L. (1985). Emerging paradigms in organizational theory and research. In Y. Lippin, (Ed.), Organizational theory and Inquiry (pp. 137-160). Beverly Hills CA: Sage.
- Coleman, P. (1985). The good school district-a critical examination of the adequacy of student achievement and per pupil expenditures as measures of school district effectiveness. Unpublished manuscript (B.C. Ministry of Education files).
- Conley, D. T. (1988). District performance standards: Missing link for effective teacher evaluation. NASSP Bulletin. November, pp. 78-83.
- Cooley, W. W., & Lohnes, P. R. (1976). Evaluation Research in Education. Irvington Publishers.
- Coombs, C. (1988). Toward a conception of educational leadership. Unpublished doctoral dissertation, University of British Columbia, Vancouver.
- Cronbach, L. (1983). Course improvement through evaluation. In G. F. Madeus, M. S. Scriven, & D.L. Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation (pp. 101-116). Boston: Kluwer-Nijhoff.
- Cronbach, L. (1982). Designing evaluations of educational and social programs. San Francisco: Jossey-Bass.
- David, J. L. (1988). The use of indicators by school districts: Aid or threat to improvement? Phi Delta Kappan. 69(7), 499-502.
- Deal, T. E. & Kennedy, A. (1982). Corporate Cultures. Reading, Mass.: Addison-Wesley.

- Dickson, G. (1988a). [The information-evaluation project: Pilot study evaluation]. Unpublished raw data.
- Dickson, G. (1988b). [Interpreting data for evaluation purposes]. Unpublished raw data.
- Doig, J. W. & Hargrove, E. C. (Eds.). (1987). Leadership and Innovation: A biographical perspective on entrepreneurs in government. Baltimore: John Hopkins Press.
- Doyle, D., & Levine, M. (1986). American business and public education: The question of quality. Peabody Journal of Education, 63(2), 17-26.
- British Columbia Teachers' Federation. (1988). Discussion guide on the Royal Commission on education. Vancouver, BC: British Columbia Teachers' Federation.
- Dunsire, A. (1973). Administration, the word and the science. London: Martin Robertson and Halstead Press.
- Eisner, E. W. (1979). The educational imagination. New York: MacMillan Publishing.
- Etzioni, A. (1980). Compliance Structures. In A. Etzioni & E. W. Lehman, (Eds.), A sociological Reader on complex organizations (third ed.; pp. 87-100). New York: Holt, Rinehart & Winston.
- Etzioni, A. (1964). Modern Organizations. Englewood Cliffs, NJ: Prentice-Hall.
- Farley, J. (1987). Justifying conclusions in naturalistic evaluations: An interpretive perspective. Evaluation and Program Planning, 10(4), 343-350.
- Fleming, T. (1988). In the imperial age and after: Patterns of British Columbia school leadership and the institution of the superintendency 1849-1988. Unpublished manuscript, University of Victoria, Department of Psychology and Social Foundations, Victoria.
- Floden, R. E., Porter, A. C., Alford, L. E., Freeman, D. J., Irwin, S., Schmidt, W. H., & Schwille, J. R. (1988). Instructional leadership at the district level: A closer look at autonomy and control. Educational Administration Quarterly, 24(2), 96-124.
- Floden, R., & Weiner, S. (1983). Rationality to ritual: The multiple roles of evaluation in governmental process. In G. F. Madeus, M. S. Scriven, & D.L. Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation (pp. 177-188). Boston: Kluwer-Nijhoff.

- Ford, J. D. (1981). Departmental context and formal structure as constraints on leader behavior. Academy of Management Journal, 24(2), 274-288.
- Freeston, K. (1987). Leader substitutes in educational organizations. Educational Administration Quarterly, 23(2), 45-60.
- Gamoran, A., & Dreeben, R. (1986). Coupling and control in educational organizations. Administrative Science Quarterly, 31(4), 49-71.
- Getzels, J. W., & Guba, E. G. (1957). Social behavior and the administrative process. School Review, 65, 423-441.
- Glasman, N.S. (1986). Evaluation-based leadership. Albany: State University of New York Press.
- Green, J. (1970). Introduction to measurement and evaluation. New York: Mead & Co.
- Greene, D., & David, J. L. (1984). A research design for generalizing from multiple case studies. Evaluation and Program Planning, 7, 73-85.
- Greene, J., Doughty, J., Marquart, J. M., Ray, M. L., & Roberts, L. (1988). Qualitative Evaluation Audits in Practice. Evaluation Review, 12(4), 352-375.
- Greene, J. (1987a). Stakeholder participation in evaluation design: Is it worth the effort? Evaluation and Program Planning, 10(4), 379-394.
- Greene, J. (1987b). Justifying conclusions in naturalistic evaluations: A practical perspective. Evaluation and Program Planning, 10(4), 325-334.
- Greenfield, T. B. (1975). Theory about organizations: A new perspective and its implications for schools. In P. Gronn (Ed.), T. B. Greenfield & his critics (pp. 63-87). Victoria, AUS: Deakin University Press. (Reprinted from Administering education: International challenge, 1975, London: Athlone Press. pp 71-99).
- Greenfield, T. B. (1979). Organization theory as ideology. In P. Gronn (Ed.), T. B. Greenfield & his critics (pp. 46-62). Victoria, AUS: Deakin University Press. (Reprinted from Curriculum Inquiry, 1979, 9(2), 97-112).
- Greenfield, T. B. (1980). The man who comes back through the door in the wall: discovering truth, discovering self, discovering organizations. In P. Gronn (Ed.), T. B. Greenfield & his critics (pp. 88-118). Victoria, AUS: Deakin University Press. (Reprinted from Educational Administration Quarterly, 1980, 16(3), 25-59).

- Griffiths, D. E. (1983). Intellectual turmoil in educational administration. In P. Gronn (Ed.), T. B. Greenfield & his critics (pp. 88-118). Victoria, AUS: Deakin University Press. Victoria, AUS: Deakin University Press. (Reprinted from Educational Administration Quarterly, 1979, 15(3), 43-65).
- Guba, E., & Lincoln, Y. (1983). Epistemological and methodological bases of naturalistic inquiry. In G. F. Madeus, M. S. Scriven, & D.L. Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation (pp. 311-334). Boston: Kluwer-Nijhoff.
- Halpin, A. W., & Winer, B.J. (1957). A factorial study of the leader behavior descriptions. In R. M. Stogdill & A. E. Coons (Eds.), Leader behavior: Its description and measurement (pp. 39-51). Columbus: Bureau of Business Research, Ohio State University.
- Hanson, E. M. (1985). Educational administration and organizational behavior (2nd ed.). Newton, Mass: Allyn & Bacon.
- Hanushek, E. (1986). The economics of schooling: Production and efficiency in public schools. Journal of Economic Literature, 24, 1141-1177.
- Havelock, R. G. (1973). The change agent's guide to innovation in education. City, NJ: Educational Technology Publications.
- Hiley, D. R. (1987). Power and values in corporate life. Journal of Business Ethics, 6, 343-353.
- Hodges, H. (1982). Madison Prep--alternatives through learning styles. In Student Learning Styles and Brain Behavior, Selected Papers from the National Conference of the Learning Styles Network (pp. 28-32). Reston, Virginia: National Association of Secondary School Principals.
- Hodgkinson, C. (1978). Towards a philosophy of administration. Oxford: Basil Blackwell.
- Hodgkinson, C. (1983). The philosophy of leadership. Oxford: Basil Blackwell.
- Holmes, M. (1986). Comment. Interchange, 17(2), 89-90.
- House, E. R. (1983). Assumptions underlying evaluation models. In G. F. Madeus, M. S. Scriven, & D.L. Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation (pp. 45-64). Boston: Kluwer-Nijhoff.

- House, E. R. (1983). How we think about evaluation. In E. House, (Ed.), How we think about evaluation (pp. 5-25). San Francisco: Jossey-Bass.
- Howell, J. P., & Dorfman, P. W. (1981). Substitutes for leadership: Test of a construct. Academy of Management Journal, 24, 714-728.
- in 't Veld, R. J., Spee, A. A. J., Tseng, H. S. S., & Sandbergen, S. (1987). Performance indicators in higher education in the Netherlands with a digression on primary and secondary education. Presented at the International Conference on Education Indicators, Washington, DC.
- Karr, A. (1987). Research backgrounder on planned change in education. Victoria, BC: Program Evaluation Branch, Ministry of Education.
- Kemmis, S., & McTaggart, R. (1988). Introduction: The nature of action research. In S. Kemmis & R. McTaggart (Eds.), The action research planner (revised ed.). Victoria, AUS: Deakin University.
- Kerr, S., & Jermier, J. M. (1978). Substitutes for leadership: Their meaning and measurement. Organizational behavior and human performance, 22, 375-403.
- Johnston, H. J. (1987). Values, culture and the effective school. NASSP Bulletin, 71(499), 79-88.
- Johnstone, J. N. (1981). Indicators of education systems. London: Kegan Page.
- Joint Committee on Standards for Educational Evaluation. (1981). Standards for evaluations of educational programs, projects, and materials. New York: Kinsport Press.
- Lecompte, M. D., & Goetz, J. P. (1984). Ethnographic data collection in evaluation research. In D. M. Fetterman (Ed.), Ethnography in Educational Evaluation (pp. 37-59). Beverly Hills, CA: Sage.
- Legislative Staff. (1987, September 19). Cash starved school system can't do it all. Times Colonist, p.5.
- Liethwood, K., & Montgomery, D. J. (1986). Improving principal effectiveness: The principal profile. Toronto: OISE Press.
- Lincoln, Y. (1985). The substance of the emergent paradigm: Implications for researchers. In Y. Lincoln, (Ed.), Organizational theory and Inquiry (pp. 137-160). Beverly Hills CA: Sage.
- Lincoln, Y. S., & Guba, E. (1985). Naturalistic inquiry. Beverly Hills: Sage.

- MacClintock, C. (1987). Administrators as information brokers: A managerial perspective on naturalistic evaluation. Evaluation and Program Planning, 10(4), 315-324.
- MacGregor, D. 1960. The human side of enterprise. New York: McGraw Hill.
- McPherson, R. B., Crowson, R. L., & Pitner, N. J. (1986). Managing uncertainty. Columbus: Charles E. Merrill.
- McIver, S. (1985). Schools planning report to the executive committee: Report #1. Draft paper presented to the executive committee of the Ministry of Education, Victoria, BC.
- March, J. G. (1981). Decision making perspective: Decisions in organizations and theories of choice. In A. Van de Ven & W. Joyce (Eds.) Perspectives on organization design and behavior. New York: John Wiley.
- Morgan, G. 1986. Images of Organization. Beverly Hills, CA: Sage.
- Moran, K. T. 1987. Research and managerial strategies for integrating evaluation research into agency decision making. Evaluation Review, II(5), 612-630.
- Murnane, R. J. (1987). Improving education indicators and economic indicators: The same problems? Educational Evaluation and Policy Analysis, 9(2), 101-116.
- Murnane, R. J., & Pauly, E. W. (1988). Lessons from comparing educational and economic indicators. Phi Delta Kappan, 69(7), 509-513.
- Murphy, P. (1987, September 24). Instruction only campaign scrubbed. Times Colonist, p. 18.
- Mussio, J. (1985). Evaluation/information project proposed policies. Draft report presented to a meeting of the Executive Committee, Victoria, BC.
- Mussio, J. (1987). Evaluation project: Follow-up strategies. Memorandum presented at a meeting of the B.C. Ministry of Education Executive Committee, Victoria, BC.
- Nechako, (1987a). The Nechako report.(in-house) Victoria: Manning .
- Nechako, (1987b). The Nechako report. (public document) Victoria: Manning.
- Nilsson, N., & Hogben, D. (1983). Metaevaluation. In E. House, (Ed.), How we think about evaluation (pp. 5-25). San Francisco: Jossey-Bass.

- Parsons, Tony. (1987, September 16). BCTF/work to rule. BCTV 6:00 News Hour. Vancouver, B.C.
- Patton, M. Q. (1987). Creative evaluation. Beverly Hills, CA: Sage.
- Patton, M. Q. (1980). Qualitative evaluation methods. Beverly Hills, CA: Sage.
- Patton, M. Q. (1978). Utilization-focused evaluation. Beverly Hills, CA: Sage.
- Pearsol, J. (1987). Justifying conclusions in naturalistic evaluations: A practical perspective. Evaluation and Program Planning, 10(4), 335-342.
- Pitner, N. J. (1986). Substitutes for principal leader behavior: An exploratory study. Educational Administration Quarterly, 22(2), 23-43.
- Policy and Planning Department. (1986). Evaluation project. Draft report, #2, presented to a meeting of the Executive Committee, Victoria, BC: BC Ministry of Education.
- Porter, A. (1988). Indicators: Objective data or political tool? Phi Delta Kappan. 69(7), 503-508.
- Program Evaluation Branch. (1988a). Five year plan. Document presented and approved by the Executive Committee, Victoria, BC: BC Ministry of Education.
- Program Evaluation Branch. (1988b). The information profile system. A paper published to explain the Information Profile System to educational professionals in BC. Victoria, BC: BC Ministry of Education.
- Program Evaluation Branch. (1988c). District information sheets. (Available from Program Evaluation & Research Branch, BC Ministry of Education, 617 Government Street, Victoria BC).
- Richards, C. E. (1988). Indicators and three types of educational monitoring systems: Implications for design. Phi Delta Kappan. 69(7), 495-498.
- Rossi, R. J., & Gilmartin, K. J. (1980). The handbook of social indicators: Sources, characteristics and analysis. New York: Garland.
- Roethlisberger, F. J., & Dickson, W.J. (1939). Management and the worker. Cambridge, Mass.: Harvard University Press.

- Royal Commission on Education: Summary of Report. (1988). A legacy for learners: summary of findings. Victoria: Queen's Printer.
- Scriven, M. (1983a). The Evaluation Taboo. In E. House, (Ed.), How we think about evaluation (pp. 75-82). San Francisco: Jossey Bass.
- Scriven, M. (1983b). Evaluation Ideologies. In G. F. Madeus, M. S. Scriven, & D.L. Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation (pp. 229-260). Boston: Kluwer-Nijhoff.
- Scriven, M. (1974). Evaluation perspectives and procedures. In W. J. Popham (Ed.), Evaluation in education (pp. 1-94). Berkeley, CA: McCutcheon.
- Sergiovanni, T. J., & Corbally, J. E. (1984). Leadership and organizational culture: New perspectives on administrative theory and practice. Urbana: University of Illinois Press.
- Sergiovanni, T. J., Burlingame, M., Coombs, F. S., & Thurston, P. W. (1987). Educational governance and administration (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Shapiro, J. P. (1988). Participatory evaluation: Towards a transformation of assessment for women's studies programs and projects. Educational Evaluation and Policy Analysis, 10(3), 191-199.
- Smith, B. (1988, March 20). McMurphy Predicts Strikes. Vancouver Sun, p. A3.
- Smith, J. K. (1983). Quantitative versus interpretive: The problem of conducting social inquiry. In E. House, (Ed.), How we think about evaluation (pp. 27-52). San Francisco: Jossey-Bass.
- Smith, J. K. (1987). Toward the justification of claims in evaluation research. Evaluation and Program Planning, 10(4), 309-314.
- Smith, J. K., & Heshusius, L. (1986). Closing down the conversation: The end of the quantitative-qualitative debate among educational inquirers. Educational Researcher, 4-12.
- Spradley, J. P. (1979). The ethnographic interview. New York: Holt, Rinehart & Winston.
- Stake, R. (1983). The case study method in social inquiry. In G. F. Madeus, M. S. Scriven, & D.L. Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation (pp. 279-286). Boston: Kluwer-Nijhoff.
- Storey, V. (1987). Leadership in uncertain times. Victoria: Research Report to the BC Ministry of Education.

- Stoney, V. (1988). Principals of Tomorrow. Vancouver: Eduserve & the British Columbia Principals' and Vice-Principals' Association.
- Stout, R. (1986). Executive action and values. Issues in Education, 4(3), 198-214.
- Stufflebeam, D. (1989, March 8). The state of the art in educational evaluation: Implementing professional standards of good practice. Paper presented at a meeting of the Washington Educational Research Association, Seattle, WA.
- Stufflebeam, D. (1983). The CIPP model for program evaluation. In G. F. Madeus, M. S. Scriven, & D.L. Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation (pp. 117-142). Boston: Kluwer-Nijhoff.
- Stufflebeam, D. L., & Welch, W. L. (1986). Review of research on program evaluation in United States school districts. Educational Administration Quarterly, 22(3), 150-170.
- Stufflebeam, D. L., & Webster, W. J. (1988). Evaluation as an administrative function. In N. J. Boyan (Ed.), Handbook of research on educational administration. Longman: New York.
- Stufflebeam, D. and Webster, W. (1983). An analysis of alternative approaches to evaluation. In G. F. Madeus, M. S. Scriven, & D.L. Stufflebeam (Eds.), Evaluation models: viewpoints on educational and human services evaluation. Boston: Kluwer-Nijhoff.
- Teddle, C., Kirby, P. C., & Stringfield, S. (1989). Effective versus ineffective schools: Observable differences in the classroom. American Journal of Education, May, 221-236.
- Webster's New Collegiate Dictionary. (1977). Springfield, MA: G. & C. Merriam Co.
- Weick, K. (1985). Sources of order in underorganized systems: Themes in recent organizational theory. In Y. S. Lincoln (Ed.), Organizational theory and Inquiry. Beverly Hills, CA: Sage.
- Whitmore, E. & Ray, M. L. (1989). Qualitative evaluation audits: Continuation of the discussion. Evaluation Review, 13(1), 78- 90.
- Wholey, J. S. (1983). Evaluation and effective public management. Boston: Little, Brown and Company.
- Willower, D. J. (1983). Contemporary issues in theory in educational administration. In P. Gronn (Ed.), T. B. Greenfield & his critics (pp. 140-162). Victoria, AUS: Deakin University Press. (Reprinted from Educational Administration Quarterly, 1980, 16(3), 1-25).

- Wirt, F. M., & Christovich, L. (1989). Administrators' perceptions of policy influence: Conflict management styles and roles. Education Administration Quarterly, XXV(1), 5-36.
- U.S. Department of Education. (1987). Elementary and secondary education indicators in brief. Washington: US Gov. Printing Office.
- Wright, D. Towards a theory of moral development. Culture, Education & Society, 40, 419-424.
- Zuckert, C. (1988). On the role of spiritedness in politics. In C. Zuckert, (Ed.), Understanding the political spirit: Philosophical investigations from Socrates to Nietzsche (pp. 1-29). New Haven: Yale University Press.

APPENDIX A1

Appendix A1: The Four-Goal and Six Attribute Framework**DRAFT COPY****INTELLECTUAL DEVELOPMENT**

-
- "To develop the ability of individuals to analyze critically, to reason and think independently, and to acquire basic learning skills and bodies of knowledge.
 - To develop in individuals a lifelong appreciation of learning, a curiosity about the world around them, and a capacity for creative thought and expression."
-

Individuals should develop the ability:

- to acquire basic learning skills and bodies of knowledge;
- to reason and think independently; and
- to analyze critically.

Individuals should also develop:

- a lifelong appreciation of learning;
- a curiosity about the world; and
- creative thought and expression.

VOCATIONAL DEVELOPMENT

-
- "To assist in developing in individuals the means to realize their personal potential and to attain their occupational and career objectives.
 - To assist in developing in individuals effective work habits, a regard for the dignity of work, and the flexibility to deal with change and challenge in the work place
 - To assist in developing in individuals the sense of self-reliance, vision, and confidence vital to the creation of new opportunities for employment."
-

The school system should assist students in developing the skills and abilities necessary to:

- reach their occupational and career objectives;
- develop effective work habits; and
- deal with change and challenge in the work place.

The school system should also assist students in developing such personal qualities as:

- self-discipline, reliability and confidence; and
- teamwork and respect for others.

Appendix A1 (cont.): The Four-Goal and Six Attribute Framework

SOCIAL DEVELOPMENT

- "To assist in developing in individuals an understanding concerning their relationship to the world in which they live, the role of the family in society, and their responsibilities and rights as citizens.
 - To assist in developing in individuals a sense of social responsibility, a tolerance and respect for the ideas and beliefs of others, and an ability to live and work with other people.
 - To assist in developing in individuals an appreciation for the fine arts and an awareness of the cultural heritage and values of Canadian society and broader world."
-

The home, the school, and the community all contribute to a student's social development. A socially developed student will:

- be knowledgeable about the world;
- be knowledgeable about the Canadian culture;
- respect the laws of the Canadian society;
- act as a responsible citizen;
- tolerate and respect the ideas and beliefs of other people and other cultures; and
- appreciate the fine arts.

HUMAN DEVELOPMENT

- "To assist in developing in individuals a sense of self-worth, individualism, personal initiative, self-discipline, the importance of the pursuit of excellence, and the satisfaction gained through achievement.
 - To assist in developing in individuals an understanding of the importance of physical health, emotional well-being and the work and leisure activities which contribute to a good quality of life."
-

A student who is well developed in this goal will exhibit traits which suggest:

- self-worth;
- personal initiative;
- discipline, including attention to physical and emotional well-being;
- appreciation of the value of leisure activities; and
- commitment to quality performance.

Appendix A1 (cont.): The Four-Goal and Six Attribute FrameworkACCESSIBILITY

-
- "Universality is vital to the education provided to B.C. children. It is not enough to say that places are provided in programs throughout the Province. Real, rather than nominal, access means paying attention to the special needs of rural, gifted, handicapped, female and minority language students. Attention does not mean conceding standards but rather exploring new methods of delivery and communication in the interests of accessibility."
-

The education system should provide as equal as possible an opportunity for all students to attend school and take all potentially necessary courses.

Access or location should not be restricted by:

- geographic location;
- ethnic, social/economic, or physical status; or
- size of community.

RELEVANCE

-
- "Excellence alone is not enough if the material taught is not current enough to catch the imagination of students and to relate what they learn through schooling to the real world as it exists now."
-

The Curriculum and teaching methods need to:

- appeal to students;
- meet basic expectations of society;
- reflect current technology;
- be adaptable to the future; and
- be reviewed regularly.

PROFESSIONALISM

-
- "Quality schooling depends on the quality of its teaching force. Continuing development of teaching skills will be even more important in the years to come as teachers will have to adjust to changing student composition (e.g. more primary students) and interests (e.g. more secondary students taking mathematics and sciences). Moreover, the role model that teachers provide in the classroom and in all of the other activities they are involved in is absolutely essential."
-

A professional teaching force should:

- be well qualified, in terms of education and experience;
- recognize the diversity of students' needs;
- continue to develop their professional skills;
- be informed about current teaching methods; and
- be positive role models for students.

Appendix A1 (cont.): The Four-Goal and Six Attribute Framework**MANAGEMENT AND ACCOUNTABILITY**

-
- "The best and most well-meaning of organizations can lose their way and fail to achieve important standards if they become preoccupied with their own existence, do not monitor and assess achievements and do not account to the public. It is absolutely vital that schooling be monitored closely and the public informed on a regular basis of the strengths and weaknesses."
-

A well-managed school system will:

- monitor student performance;
- put information about student performance to use;
- take initiatives to enhance student achievement;
- answer to the student body, taxpayers, parents and public;
- share information on resource utilization and achievement with those groups;
- make the link between inputs (e.g.: proposed tax increases) and expected outcomes (e.g.: improved performance on learning assessment exams);
- know what they are attempting to achieve; and
- strive to make the most effective use of available resources.

COST EFFECTIVENESS

-
- "In a diverse society with an aging population, schooling is only one of many social responsibilities undertaken by government. Respecting the need to balance all social responsibilities and to make them all cost-effective, is a vital component of responsible government in an increasingly competitive world economy."
-

A cost effective system will:

- maximize the return on resources invested;
- achieve the highest level of efficiency where possible;
- relate system and student outputs (effectiveness) to resources used (cost); and
- review expenditures and financial allocations in light of their mission statement, objectives, and goals.

PUBLIC EXPECTATIONS

-
- "In a democratic society, it is not enough to provide the best possible schooling. That schooling must be provided so that it helps to define and responds to public expectations."
-

System managers must attempt to:

- identify public expectations;
- insure proper dissemination of information for public consideration;
- attempt to meet those concerns in its allocation of human and financial resources;
- show the results of strategic planning; and
- ensure local schools reflect community needs and interests.

Appendix A2: Sample pages of an Information Profile From a B.C. School District.

RUN 88/02/16

MINISTRY OF EDUCATION - 1987/88 SCHOOL DISTRICT INFORMATION PROFILES

PAGE 7

District 085 VANCOUVER ISLAND NORTH

1.0B Intellectual Development (GRADE 12)
1.2B Provincial Exam Results

Note: No. of students and participation rates based on Grade 12 population only

ENGLISH LIT 12

	Participation Rate (%)	No. of Students	Exam Mean	School Mean	A/B	Final Results (%) C+/C/P	F
District	9.1	15	63	64	40.0	40.0	20.0
Category D5	7.6	117	55	66	26.5	53.8	19.7
Province	11.9	3,872	64	68	34.6	58.3	7.0

ENGLISH 12

	Participation Rate (%)	No. of Students	Exam Mean	School Mean	A/B	Final Results (%) C+/C/P	F
District	73.9	122	65	66	25.4	71.3	3.3
Category D5	73.8	1,131	64	67	26.9	68.3	4.9
Province	77.4	25,134	67	67	31.9	63.9	4.2

FRENCH 12

	Participation Rate (%)	No. of Students	Exam Mean	School Mean	A/B	Final Results (%) C+/C/P	F
District	15.8	26	58	70	23.1	65.4	11.5
Category D5	9.7	149	62	71	29.5	67.1	3.4
Province	14.3	4,646	67	71	42.5	53.1	4.4

GEOGRAPHY 12

	Participation Rate (%)	No. of Students	Exam Mean	School Mean	A/B	Final Results (%) C+/C/P	F
District	18.8	31	57	61	9.7	71.0	19.4
Category D5	27.5	422	65	66	27.3	66.8	5.9
Province	25.3	8,227	65	65	26.6	67.4	6.0

RUN 88/03/18

MINISTRY OF EDUCATION - 1987/88 SCHOOL DISTRICT INFORMATION PROFILES

PAGE 18

District 085

VANCOUVER ISLAND NORTH

7.0	Cost Effectiveness
7.2	Budgeted Costs

1987/1988		Fcn 1	Fcn 2	Fcn 3	Fcn 4	Fcn 5	Fcn 6	FCN 7
District	Budget	4,559,632	3,121,698	1,117,960	737,869	2,329,691	144,421	692,859
	% 1-7	35.89	24.57	8.80	5.81	18.34	1.14	5.45
	CPP	(FCNS 1, 2 AND 3 COMBINED)		3,044	255	806	50	240
Category D5	Budget	40,024,908	30,460,723	11,495,473	5,593,996	19,571,319	950,536	6,517,952
	% 1-7	34.92	26.58	10.03	4.88	17.08	0.83	5.69
	CPP	(FCNS 1, 2 AND 3 COMBINED)		2,912	198	694	32	230
Province	Budget	635,058,453	511,897,949	189,100,572	83,044,473	282,429,250	35,360,916	51,840,361
	% 1-7	35.50	28.62	10.57	4.64	15.79	1.98	2.90
	CPP	(FCNS 1, 2 AND 3 COMBINED)		2,863	178	605	76	111
1986/1987								
District	Budget	4,373,632	3,200,186	913,777	631,902	2,119,972	124,345	652,694
	% 1-7	36.40	26.63	7.60	5.26	17.64	1.03	5.43
	CPP	(FCNS 1, 2 AND 3 COMBINED)		2,897	216	724	42	223
Category D5	Budget	37,954,833	29,271,529	10,352,829	5,184,035	18,837,078	910,117	5,972,536
	% 1-7	34.99	26.98	9.54	4.78	17.36	0.84	5.51
	CPP	(FCNS 1, 2 AND 3 COMBINED)		2,709	179	656	30	206
Province	Budget	587,231,413	487,507,347	166,538,002	75,736,454	262,523,304	33,710,433	48,043,716
	% 1-7	35.35	29.35	10.02	4.56	15.80	2.04	2.89
	CPP	(FCNS 1, 2 AND 3 COMBINED)		2,667	163	564	72	103
1985/1986								
District	Budget	4,440,299	3,163,145	847,643	635,996	2,080,241	138,627	652,960
	% 1-7	37.13	26.45	7.09	5.32	17.39	1.16	5.46
	CPP	(FCNS 1, 2 AND 3 COMBINED)		2,826	213	696	45	218
Category D5	Budget	36,956,833	28,983,556	9,161,171	4,949,251	18,002,527	894,486	5,892,862
	% 1-7	35.25	27.65	8.74	4.72	17.17	0.85	5.62
	CPP	(FCNS 1, 2 AND 3 COMBINED)		2,585	168	618	29	201
Province	Budget	556,858,786	465,473,319	146,548,496	71,906,505	245,228,021	32,644,331	45,228,545
	% 1-7	35.61	29.76	9.37	4.60	15.68	2.09	2.89
	CPP	(FCNS 1, 2 AND 3 COMBINED)		2,499	154	521	70	97

Note: Costs per pupil (CPP) are calculated by dividing costs (such as gross operating budget) by the appropriate full time equivalent (FTE) pupils.

Appendix A3 (cont.): Sample of one of the seven Questionnaires used in the Inlet and South Coast District evaluations.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45

Indicate if you agree or disagree with the following statements.

- 5. Teachers in this school care about me as an individual.....
- 6. The staff and students were positive and supportive of each other.....
- 7. School provided me with the knowledge, skills and attitudes required for today's world.....
- 8. The skills and knowledge taught at school helped me obtain a job.....
- 9. What was taught at school prepared me for college and university.....
- 10. The school provided help for applications to colleges and universities....
- 11. The school administration exercised appropriate leadership.....
- 12. The School District offered the right programs of studies.....
- 13. The School District provided the proper leadership.....
- 14. The School Board (Board of Trustees) provided proper leadership.....

Part B: Information About You.

- 15. Did you belong to a school club?
- 16. Did you play intramural or house games at noon or after school?
- 17. Did you belong to a school team?
- 18. Did you help out with school activities?
- 19. Did you receive a school leaving certificate, but not Graduate?
- 20. Did you officially graduate from school?
- 21. What sex are you? Male Female
- 22. Your grade average in your last year was A B C± D F
- 23. What are you currently doing? (fill in all that apply)

<input type="radio"/> Studying <input type="radio"/> by correspondence <input type="radio"/> at a secondary school <input type="radio"/> at a post-secondary institute	<input type="radio"/> Working <input type="radio"/> full time employment <input type="radio"/> part-time employment <input type="radio"/> unemployed
---	---

THANK YOU VERY MUCH FOR YOUR HELP IN COMPLETING THIS QUESTIONNAIRE.
(Comments are welcomed in the comment area on the back of the front cover.)

CONFIDENTIALITY ASSURED

A	B	C	D	E	F	G	H	I	J
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1	2	3	4	5	6	7	8	9	10

FOR OFFICE USE ONLY

APPENDIX B1

Appendix B1: Interior District Interview Protocol.***Introduction and explanation of interview:***

1. The *process* used to develop information for the report, and to develop the report itself has not traditionally been used by evaluators at the school district level.

The evaluation team was given great leeway to tell the evaluator (and the school board) what information was valuable, what information was meaningful, how to interpret this information, and how to present it in the form of a report. Usually the evaluator is given the task of evaluating a program and then proceeds independently to do so: but the evaluation team formulated in the district completed much of these tasks themselves.

2. My interest and the reason for the interview is primarily to find out about the potential use of this form of evaluation--in particular, as a leadership tool in the school district.

Such an instrument must accomplish leadership goals, be applicable to a number of different kinds of districts, while at the same time be reflective of their unique needs and character. In addition it must capture the reality of education, in all its many facets, accurately and honestly.

3. I personally feel there were both benefits and problems with this kind of evaluation process. You'll have a chance to talk about both in this interview. Please be as honest and complete in our discussion as you possibly can.

4. Please try to elaborate on the reasons why you answered a particular question when possible.

Confidentiality

1. I will need to take notes during our talk, so I can remember what is said. But, I want to assure you that our discussion will remain confidential, in that your *name* will *not* be used when I write my paper. I will:

- a. summarize what a number of interviewees say, and;
- b. use selected quotes, but not attach your name to them.

2. The Ministry or the school board will not be given specific data from the interviews, only the paper, which will give summarized results.

Interview Format

1. This will be a semi-formal interview, but we'll try to make it as informal as possible.

2. Our discussion should take between 45 minutes and an hour.

3. I appreciate your willingness to help me out more than you already have on this project.

Appendix B1 (cont.): Interior District Interview Protocol, p.2.

Evaluation Project Interview

1. School district

2. Interviewee:

Representative (of which stakeholder group) Title:

3. Interviewer:

4. Describe general tone of the interview:

5. Other comments regarding interview:

Appendix B1 (cont.): Interior District Interview Protocol, p.3.

Question One: Let's discuss your role as a member of the evaluation team.

a. How would you characterize or describe your role during the process of evaluation?

I'm going to give you a set of cards with words on them, words which may relate to your view of the role you played in the evaluation process. Please read through these words, then separate the cards into three stacks:

Describe role **Well** Describe role **Somewhat** Do **Not** describe role

Which **three** words on these cards **best** describe your role?

Why?

Appendix B1 (cont.): Interior District Interview Protocol, p. 4.

partnership	director	subordinate
observer	controller	active
inactive	opinionated	reticent
positive	negative	knowledgeable
ignorant	trusted	threatened
cooperating	managing	acquiescing
shy	involved	uninvolved
passive	decisionmaker	participant
follower	collaborator	yielding
powerless	powerful	compliant
strong	weak	indulged
important	insignificant	appeasing

2. Which *three* words best describe the attitude and behaviors of the evaluation team as a whole?

Why?

Appendix B1 (cont.): Interior District Interview Protocol, p.5.

3. a. What is your general *opinion* of the *process* used to gather information for the report

b. Please describe any positive aspects or features of the process we used to develop the report.

c. Please describe any negative aspects or features of the process we used to develop the report.

d. Should the overall process be used regularly and systematically for system improvement?

Appendix B1 (cont.): Interior District Interview Protocol, p.6.

4. The evaluation team used a number of specific procedures to organize and gather information from the various groups involved in education in the school district. As I name each method, please fill in the two cards I'll give you to rate each procedures used.

Card 4:

1. The Use of the Four Goal/Six Attribute Framework

positive	1	2	3	4	5	6	7	negative
accurate	1	2	3	4	5	6	7	faulty

2. The development and use of district-based key indicators of educational performance

positive	1	2	3	4	5	6	7	negative
accurate	1	2	3	4	5	6	7	faulty

3. The use of government exam data, retention rates, learning assessment performance of students as provided by Ministry "key indicators"

positive	1	2	3	4	5	6	7	negative
accurate	1	2	3	4	5	6	7	faulty

4. The use of surveys of various constituents (parents, non-parents, etc.)

positive	1	2	3	4	5	6	7	negative
accurate	1	2	3	4	5	6	7	faulty

5. The use of a report designed to report to the management team and stakeholder groups (detailed report)

positive	1	2	3	4	5	6	7	negative
accurate	1	2	3	4	5	6	7	faulty

6. The use of a report designed to report to the general public (parents, non-parents, and employers, etc.).

positive	1	2	3	4	5	6	7	negative
accurate	1	2	3	4	5	6	7	faulty

7. The use of the Q-sort to decrease the number of indicators chosen.

positive	1	2	3	4	5	6	7	negative
accurate	1	2	3	4	5	6	7	faulty

8. The use of a broad based stakeholder team

positive	1	2	3	4	5	6	7	negative
accurate	1	2	3	4	5	6	7	faulty

Appendix B1 (cont.): Interior District Interview Protocol, p.7.**Card 5:**

Here is a coded scale, from 1 to 7 expressing a continuum between a positive characteristic and a negative characteristic. Applying these to the evaluation process, and procedures, give your evaluation of its success for each pair of terms:

Just	1	2	3	4	5	6	7	Unjust
Open	1	2	3	4	5	6	7	Closed
Non-Educative	1	2	3	4	5	6	7	Educative
Irrational	1	2	3	4	5	6	7	Rational
No Integrity	1	2	3	4	5	6	7	Has Integrity
Interesting	1	2	3	4	5	6	7	Boring
Dishonest	1	2	3	4	5	6	7	Honest
Dignifying of Constituent Opinions	1	2	3	4	5	6	7	Denigrating of Constituent Opinions
Not Meaningful	1	2	3	4	5	6	7	Meaningful
Politics conducted responsibly	1	2	3	4	5	6	7	Vulnerable to Pressure Group Politics
Excessive use of executive authority	1	2	3	4	5	6	7	Minimal use of executive authority
Efficient structure	1	2	3	4	5	6	7	Confusingly complex structure
Ineffective Data Collection	1	2	3	4	5	6	7	Effective Data Collection
Functional Effectiveness	1	2	3	4	5	6	7	Duplicate Activities
Freedom within guidelines	1	2	3	4	5	6	7	Rule-bound procedures

Appendix B1 (cont.): Interior District Interview Protocol, p.8.

a. Please indicate your general, overall feeling about the evaluation process and procedures used to develop the report

Positive 1 2 3 4 5 6 7 Negative

b. Please use the scale again to indicate your overall, general feeling about the product: the report.

Positive 1 2 3 4 5 6 7 Negative

Why?

5. *Did this system allow for the uniqueness of your school district to be reflected in the evaluation process, and in the final reports?*

Why or why not?

6. *What aspects of the process and procedures would you recommend that other districts utilize if undertaking an evaluation? Why?*

Appendix B1 (cont.): Interior District Interview Protocol, p.9.

Card 6: On a seven point scale, please indicate how you agree or disagree as to whether the overall process has contributed to achievement of the following statements:

1. provides a consistent knowledge base for the management team and the general public

agree 1 2 3 4 5 6 7 disagree

2. allows for accurate determination of general strengths and weaknesses in the educational programs of the district.

agree 1 2 3 4 5 6 7 disagree

3. provides a springboard for effective short and long term planning.

agree 1 2 3 4 5 6 7 disagree

4. will develop a higher degree of shared commitment and purpose amongst professional staff in the district.

agree 1 2 3 4 5 6 7 disagree

5. will develop a higher degree of shared commitment and purpose amongst parents and general public.

agree 1 2 3 4 5 6 7 disagree

6. provides fair representation of all interest groups in education throughout the process.

agree 1 2 3 4 5 6 7 disagree

7. helps to provide a clear vision, of where the education system is and should go in the future. It considers the short and long term values of education.

agree 1 2 3 4 5 6 7 disagree

8. will improve learning for students.

agree 1 2 3 4 5 6 7 disagree

9. assesses student performance, or achievement, rather than school processes and procedures.

agree 1 2 3 4 5 6 7 disagree

Appendix B1 (cont.): Interior District Interview Protocol, p.10.

7. Please rank the 9 goals of education using the Q-sort below.

MOST IMPORTANT

LEAST
IMPORTANT

	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>

What is the reason for your top three choices.

8. Evaluation has traditionally been somewhat ineffective as a leadership tool. As a result of your experiences with this project, please rank the following list of suggested reasons on the "Q-sort" below as the most likely reasons to explain this circumstance.

1. evaluation usually does not involve stakeholders in its design.
2. evaluation is threatening because it points out areas of weakness.
3. evaluation consists of many different levels of value judgement; therefore it is difficult to maintain objectivity and rationality, and easy to manipulate results in a favorable fashion.
4. evaluation, when effective, indicates improvement needs that are impossible to achieve because follow-up is not planned well.
5. evaluation is too time-consuming and resource consuming for its worth.
6. evaluation requires an expertise that few administrators possess.

Appendix B1 (cont.): Interior District Interview Protocol, p.11.

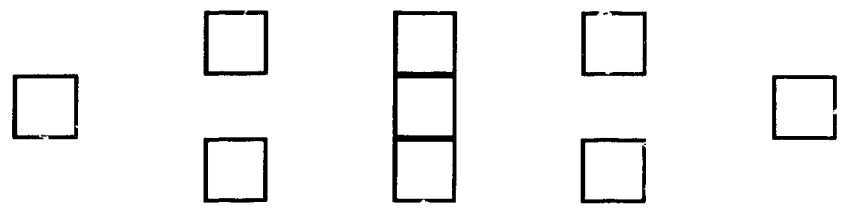
7. evaluations have been designed that do not meet the needs and requirements of local decision-makers.

8. evaluation reports are extremely difficult to write so as to report the data accurately, while at the same time providing for readability.

9. evaluation is the last activity in implementing a program, and therefore its value as a leadership exercise is often ignored.

MOST LIKELY

LEAST LIKELY



What is the reason for your top 3 choices?

9. *How could the evaluation team established better groundwork and follow up?*

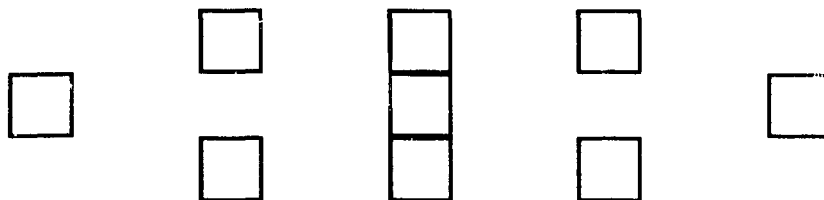
Appendix B1 (cont.): Interior District Interview Protocol, p.12.

10. To what do you attribute the success and integrity of the evaluation.

1. Senior Administrators treated committee members with respect and consideration.
2. Q-sort that allowed equal input in decision making.
3. Clearly defined tasks.
4. Alterations of conventional stakeholder roles
5. Senior administrators knowing what they wanted and persuading people to do what they wanted.
6. Novelty of process.
7. Reliance on knowledge and values of interest groups.
8. Time lines and actions that were well-structured.
9. The use of "judgement and evaluation" at each stage of decision making throughout the process.

MOST IMPORTANT

LEAST IMPORTANT



11. *Why do you think the evaluation was requested in the first place? What is your understanding of the reasons?*

Appendix B1 (cont.): Interior District Interview Protocol, p.13.

12. Do you believe that there was any relationship between the political situation in early 1988-89 and the evaluation project (either locally or provincially?) If this is the case, did it impact on the potential success of this project.

13. How would you characterize the motivations of the evaluation team members in completing this project?

14. How could better, more accurate information have been gathered for the evaluation report? Any suggestions?

Appendix B1 (cont.): Interior District Interview Protocol, p.14.

15. *Is the report a practical tool for leadership? Please Explain.*

16. *Please describe the use of the reports that have been produced; by you personally, and by the school district.*

by teachers:

extensive ___ some use ___ little ___ no use ___

by the management team:

extensive ___ some use ___ little ___ no use ___

by the public:

extensive ___ some use ___ little ___ no use ___

To what extent do you think it will be used in the future?

17. Did you come up with a summative evaluation overall about the district's performance?

18. Are you committed to district wide improvement?

Appendix B1 (cont.): Interior District Interview Protocol, p.15.

19. As a result of your participation in this project has your perception about the need for an improvement plan become stronger or weaker? Please explain.

APPENDIX B2

Appendix B2: Audit trail guide to ensure trustworthiness (adapted from Greene et al, (1988), p. 358-59.

I - Background and Initial Preparation

1. 1985 background papers
2. Pilot evaluation study
3. Ministry documentation
4. working papers during project formulation
5. Nechako reports

II - Evaluation Design Planning

1. B. Collis outline
2. Working papers re administration leadership and evaluation
3. Dissertation proposal (short) to committee
4. Initial contracts
5. IP System mailouts
6. Overheads for workshop presentation

III - Evaluation Design Implementation

1. Various drafts of proposal and design
2. Drafts of instruments
3. Ayers instrument; McPherson paper
4. Initial interviews, logs, outlines to districts.
5. Revised final design

IV - Instrument development

1. Samples of rough and final (see above)

V - Sampling

1. Districts:
 - a. notes re initial choices and why
 - b. ongoing generation (see above)
2. Respondents (interview, public, etc.):
 - a. design trail above
 - b. district choice

Appendix B.2 (cont.): Audit trail guide.

VI - Raw Data (*see samples in Appendices C, D and E).

1. Working documents: agendas, logs, etc.
2. Personal notes
3. Letters, agreements*
4. Evaluation data and procedures (ie., samples of indicators as generated*, Q-sorts* rough drafts of reports, etc.)
5. Interviews (tapes and transcripts)*
6. Questionnaires
7. Participant notes

VII - Data Summary (*see samples in Appendices C, D, E, F and G)

1. Chronology of events*
2. Individual case summation of stakeholder ratings for critique of the IPS*.
3. Individual case summation of semantic differentiation procedure*.
4. Results of Q-sort on "leader-substitutes"*

VIII - Data Synthesis and analytic notes (*see Appendix F)

1. Samples of color-coded interview cards grouped together by pattern.
2. Synthesis of stakeholder ratings of IPS over all cases*.
3. Synthesis of semantic differentiation results over all cases*.
4. Final IPS model*.
5. Summary of notes from personal dissertation journal.
6. Rough notes and rough drafts of case study and meta-evaluation chapter construction.

Appendix B.2 (cont.): Audit trail guide.Sample questions from audit guidelines

<u>Audit</u>	<u>Tasks</u>	<u>Sample Guiding Questions</u>	<u>Suggested References</u>
<i>Dependability</i>	Identify decisions and rationales, including context/purpose for method	<ol style="list-style-type: none"> 1. Is there evidence of purposive or theoretical sampling? 2. Is there adequate support for sampling decisions? 3. Are the methods used appropriate? 4. Is there adequate support for decisions about materials? 	III, IV, V
<i>Confirmability</i>	Verify linkages between data & findings	<ol style="list-style-type: none"> 1. Is there isomorphism across linkages, e.g., from raw data to analyzed data to synthesized data? 	VI, VIII, IX

APPENDIX C1

Appendix C1: Letter of agreement between Inlet School District and the Program Evaluation and Research Branch, Ministry of Education.

March 18, 1988

Mr. Bob Ney
 Superintendent of Schools
 School District #1 (Inlet)
 3981 Vancouver Street
 Cape George, B.C.

Dear Mr. Ney:

This letter will confirm fundamental aspects of the agreement between School District #1 (Inlet) and the Program Evaluation Branch of the Ministry of Education with respect to the district evaluation being done in the district this school year.

The school district will:

- * provide a "contact person" who will work with the Program Evaluation Branch facilitator and ensure that district responsibilities are carried through, such as:
 - * the gathering of questionnaire data;
 - * an analysis of district or school based data an attempt to discover "district based" indicators of educational performance;
 - * data interpretation and evaluation; and
 - * the production of a district and/or public report.
- * provide sufficient funds and personnel resources to ensure:
 - * the effective distribution and scoring of questionnaires designed to elicit public opinion, and payment of these costs (between \$5,000 and \$10,000 depending upon school district population) will be a direct responsibility of the school district; and
 - * the design, print, and distribution of either a district and/or public evaluation report.

The Program Evaluation branch will:

- * provide a facilitator to assist with the evaluation process in general, and in particular to provide consultation with respect to:
 - * the gathering of questionnaire data;
 - * data interpretation and evaluation;
 - * the production of both a district and public report based on this information; and
 - * processes and procedures that would assist the evaluative process and its follow-up improvement plans.

It is the intent of the Program Evaluation Branch to assist the school district to conduct a fair, rational, and meaningful evaluation of its educational operations.

The evaluation will be based on the framework devised by the Ministry of Education, Program Evaluation Branch (i.e., the four goals of education - intellectual, vocational, social and human development, and the six attributes of effective schools - accessibility, relevance, professionalism, management and accountability, cost-effectiveness, and public expectations).

Data for the evaluation will come from three sources:

- * the Ministry data files, through the Information Profile System;
- * questionnaires distributed to various constituents in the school district; and
- * district and school based files, as indicated by the district.

It is anticipated that questionnaire distribution will be completed by the end of the 1987-88 school year, and reporting out documents by September 30, 1988.

I trust this letter is consistent with the content and nature of our initial discussions on the topic. If so, could you indicate confirmation of your approval and acceptance of this process by signing the copy enclosed and returning it to the above address. We look forward to working with you.

Yours truly

Ian J. Cameron
Director
Program Evaluation Branch

Signing authority and title.

IC: clh
FMG: L-103

APPENDIX C2

Appendix C2: Chronology of Events in the Inlet Evaluation (*denotes meeting attended by researcher or associate).

<u>Date:</u>	<u>Task:</u>
September-December, 1987	Preliminary discussions in district
Dec. 14, 1987	Letter from Bob Ney expressing interest in evaluation project sent to Ministry.
January 30, 1988	Meeting between researcher and Bob Ney at Ministry.
March 18, 1988	Letter outlining contractual agreement sent to district.
January-April, 1988	Continued negotiations about survey in district.
*April 20, 1988	meeting with IDTA executive.
April 21-July 1, 1988	Survey administered.
*July 4, 1988	Meeting between researcher Bob Ney, and Frank Ford to establish process for two day meeting.
*July 5-6, 1988	First Stakeholder meeting: <ol style="list-style-type: none"> 1. framework acceptance 2. brainstorming 3. indicator agreement 4. survey review 5. assignment of data collection tasks
*July 19, 1988	Day of informal interviews
*July 20-21, 1988	Second Stakeholder meeting <ol style="list-style-type: none"> 1. data collating, winnowing, interpreting 2. discussion of stakeholder role in report writing
July 27-28, 1988	Writers work at Ministry.

<u>Date:</u>	<u>Task:</u>
July 29-Sept. 1988	Creation of report by graphic designer.
September 23, 1988	Stakeholder approval of report for printing
September 24-31, 1988	Printing of report
October 3, 1988	Dinner meeting with Bob Ney delivery of report.
*October 6, 1988	School Board analysis of report.
October-December, 1988	Follow-up activities.
*December 5-6, 1988	Interviews of stakeholders; Parent meeting at junior secondary.
February 1, 1989	Debrief of Bob Ney.
December, 1988-March, 1989	Continuation of follow-up activities.

APPENDIX C3

Survey Administration Procedures in the InletSchool District

The superintendent and the director made the decision to self-administer the survey--in other words, to design the sample, distribute questionnaires, and collect them on their own. The prime reason for this decision appeared to be the cost factor: total expenditures for self-administration and scoring was approximately \$1500.00. Samples were as follows:

1. Parents	200 respondents
2. Non-parents	100 respondents
3. Teachers	100 respondents
4. Secondary students	200 respondents
5. Elementary students	200 respondents
6. Ex-grads	100 respondents
7. Employers	100 respondents

Administration consisted of (1) providing school administrators with questionnaires so that teachers and students from each school in the district were proportionately represented; (2) developing a grid of the whole district, and having students deliver questionnaires to randomly selected homes in sequence, and picking same up three days later (the students were paid for this service); (3) employing a similar procedure with employers; and (5) contacting graduates (or their parents) by phone to establish where to send the graduate survey. A very high percent rate of return, ranging from 80-96%, was achieved.

APPENDIX C4

Appendix C4: Data Interpretation and Analysis Procedures in the Inlet School District Evaluation.

1. Large papers, with goal or attribute headings, were pasted around the room.
2. Stakeholder team members were asked to affix data elements under the framework headings, writing out the final key indicator definition.
3. In pairs, stakeholders cut and paste questions from the surveys, placing them under appropriate framework headings. Comparable questions from each constituent group were placed side by side.
4. Pairs were then assigned two framework elements each. They were asked to color code the information according to the following scheme:

red: necessary to report because it reflects a vital element of the framework goal being measured.

blue: a score, or measure, that appeared to represent a strength in the district.

green: a score, or measure, that appeared to represent a weakness in the district.

yellow: a score on a survey item that represented, in the stakeholder team's qualitative judgment, a significant constituent difference.

5. Each pair was then required to review their decisions with the whole group. If substantial disagreement was noted, group consensus was sought for an appropriate recoding.
6. If key indicators were judged to be inappropriately operationalized, a group decision to eliminate them from the study was made.

APPENDIX D1

Appendix D1: Chronology of Events in the South Coast Evaluation (*denotes meeting attended by researcher or associate).

<u>Date:</u>	<u>Task:</u>
January, 1988	South Coast Trustee attended Phi Delta Kappa meeting and became acquainted with evaluation project.
*Feb. 15, 1988	Meeting between Bill Nelson and researcher with respect to the participation of South Coast in an evaluation.
April 8, 1988	Letter re contractual agreement sent to district.
April 14, 1988	Signed letter returned.
February-April, 1988	Discussions at the School Board level about participation in the project.
*April 19, 1988	Initial meeting with John Lee.
April 24, 1988	Letter to steering committee members.
*May 4, 1988	First Stakeholder Collaboration Team Meeting.
May 9, 1988	Letter to Stakeholder Team members about survey assignments and responsibilities..
May 27, 1988	Letter re questionnaire sent out to teachers/administrators.
June 1, 1988	Letter re questionnaire to parents and public re questionnaire.
June 1-June 20, 1988	Administration of surveys.
*June 21, 1988	Stakeholder Team meeting to develop key indicators.

<u>Date:</u>	<u>Task:</u>
*June 28, 1988	Informal interviews
*June 28, 1988 (aft.)	Continuation of key indicator development.
June 29-August 23, 1988	Data collection on selected human and social development indicators.
*August 23, 1988	Completion of key indicator choice.
*August 25, 1988	Data synthesis, interpretation, analysis; collating and winnowing.
*September 8, 1988	Meeting between South Coast report writers and Ministry Team.
September 19, 1988	First draft of report presented to the South Coast writers and Stakeholder Team.
October 3, 1988	Second draft of report given to South Coast writers.
*October 11, 1988	School Board meeting to discuss report publication.
*October-November 21, 1988	Four meetings re report final draft preparation.
November 25, 1988	Report printed.
November, 1988-January, 1989	Production of Newspaper Insert.
February 28, 1989	Completion of Newspaper Insert
March 6-10, 1989	Distribution of Newspaper Insert during Education Week.

APPENDIX D2

Appendix D2: Survey Administration Procedures in the South Coast School District.

John Lee (the Director) made the decision to assign the responsibility of survey administration to the Stakeholder Team. There were two reasons for this: (1) the Director did not have the time to administer the survey himself; and (2) the cost of contracting out the questionnaire was too great. Responsibilities for each Team member were outlined in a letter to each person in the stakeholder group.

Response rates were as follows:

1. Parents	150/200 returned
2. Non-parents	approximately 65/100 returned
3. Teachers	80/100 returned
4. Secondary students	198/200 returned
5. Ex-grads	60-75/100 returned
6. Employers	approximately 65 of 70 returned.

Final figures as to return rates for employers and non-parents were unavailable, as Stakeholder Team members did not (or could not) keep good records of the response rate.

Administration was handled by assigning each Stakeholder Team member a specific responsibility. A letter explaining the purpose of the questionnaire was attached to each survey (see Appendix D4).

Each Stakeholder Team member utilized resources at their disposal to accomplish their tasks. Thus the parent group and district employees became involved in the survey's distribution.

APPENDIX D3

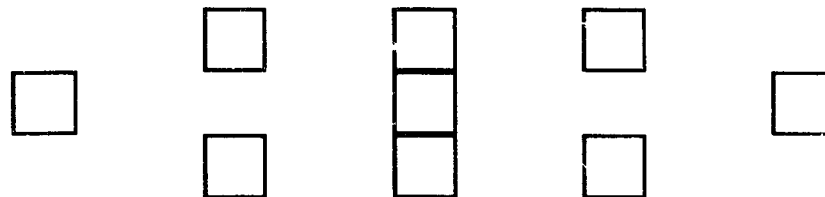
Appendix D3: The Q-sort process utilized to prioritize key indicators in the South Coast evaluation.

The Q-sort is a process that allows each stakeholder equal input into the determination of priority ranking within a list of up to nine items. It is conducted in the following manner:

A series of boxes labelled "least important" to "most important", are arranged as in Figure D3a below:

Figure D3a.

MOST LIKELY



Each item to be prioritized is then given a number. Then these steps are followed (it is important that this sequence be used in order to maintain the integrity of decisionmaking):

1. The respondent is asked to look at the list of items, and choose the one that to them is the "most important". Its number then goes in the box to the far right.

2. The respondent then looks at the remaining list and chooses the item that is "least important". The number is placed in the box in the far left.

Appendix D3 (cont.): The Q-sort process utilized to prioritize key indicators in the South Coast evaluation.

3. Then, from the remaining list, the respondent chooses two items deemed the "most important". These are placed in the two boxes to the right.

4. This step is repeated with the respondent choosing the two items from the remaining list deemed "least important". These are placed in the two boxes to the left.

5. The remaining three items are placed in the middle set of boxes.

Each item is then given a score. It is determined by giving all items in the box on the right a score of 10, the next set of boxes 8, 6, 4 and 2 accordingly. The items can then be placed in a priority list according to their score.

APPENDIX D4

Appendix D4: Letter outlining Stakeholder Team responsibilities for the survey administration in the South Coast School District.

May 9, 1988

TO: Members of the District
Assessment Committee

FROM: John Lee

Once again I would like to thank you for agreeing to be responsible for the distribution and collection of the following questionnaires:

50-70 Employees	Mrs. Jane Holden
100 Non-parents	Mrs. Shirley Eastwood and Mr. Ken Franks
20 Tunglat graduates	Mr. Dan Roebottom

I will be responsible for the following:

150 Teachers
200 Parents
200 Secondary graduates
80 Mountbatten graduates

As soon as the questionnaires are available, I will give you a call.

....2

Appendix D4 (cont.): Letter outlining Stakeholder Team responsibilities for the survey administration in the South Coast School District.

.....2

Future Meeting Dates

June 21, 1988 - 1:00 p.m. Board Room

June 28, 1988 - 1:00 p.m. Board Room

I hope that you will be able to join me for lunch on both days. Let's meet at the Carmelia at noon! If you are unable to make the lunch on either day, please give me a call.

July 18, 1988 - 9:00 a.m. to 3:30 p.m.

July 19, 1988 - 9:00 a.m. to completion of task.

cc. Mr. Graham Dickson
 Ministry of Education

APPENDIX D5

Appendix D5: Data Interpretation and Analysis Procedures in the South Coast School District Evaluation

Each person was assigned two framework elements and large papers were placed around the room for pasting and collating data. Stakeholder team members were asked to affix data elements under the framework headings. Individually stakeholders cut and paste questions from the surveys, placing them under appropriate framework headings. Comparable questions from each constituent group were placed side by side.

Stakeholder team members were asked to color code the information according to the following scheme:

pink: necessary to report because it reflects a vital element of the framework goal being measured.

blue: any survey element that received an 80% or better satisfaction rating; ie., a district strength.

green: a score, or measure, that in the judgment of the interpreter represented a below expected score; ie., a district weakness.

yellow: a score on a survey item that represented, in the individual's qualitative judgment, a significant constituent difference.

Each individual then reviewed their judgments with the remainder of the group. Disagreements were aired and final decisions on coding were made.

APPENDIX E1

Appendix E1: Chronology of Events in the Interior Evaluation(*denotes meeting attended by researcher or associate).

<u>Date:</u>	<u>Task:</u>
*July 8, 1988	Preliminary discussions in during researcher's visit to the district.
August 17, 1988	Letter from Jason Stone outlining district intent to take part in the evaluation project.
August 20-September 25, 1988	Informal discussions re survey and other procedures.
*September 27, 1988	First Stakeholder meeting in the district.
October 5, 1988	Letter outlining survey intents sent to Ministry.
*October 7, 1988	Meeting with Jason Stone to clarify project intents.
October 7-28, 1988	Jason Stone and Damon Korchuk developed list of key indicators for next stakeholder meeting.
*October 28, 1988	Second stakeholder meeting: key indicator list discussed and refined.
*November 10, 1988	Third stakeholder meeting: indicators revisited; prioritized final list.
November 10-December 14, 1988	Key indicator list refined once again; distributed to committee members.
*December 14, 1988	Meeting between Damon Korchuk and researcher to refine list of data required to measure indicators.

Appendix E1 (cont.): Chronology of Events in the Interior Evaluation.

<u>Date:</u>	<u>Task:</u>
*January 4, 1989	Fourth stakeholder meeting: Data assignments given out; meeting with survey company representative.
*January 20, 1989	Fifth Stakeholder committee meeting: Review of key indicator data; interpretation and start on report writing.
January 6-28, 1989	Survey administration.
January 29-February 24, 1989	Booklet and video writing and preparation.
*March 3, 1989	Presentation of evaluation report at British Columbia Trustees Association showcase.
March 3-April 30, 1989	Follow-up activities
*April 30, 1989	interviews

APPENDIX E2

Appendix E2: Interior School District Mission Statement

Mission Statement for Interior School District

"Our Mission is to educate students to value learning, possess knowledge and skills, strive for personal excellence, experience success, exhibit responsible citizenship, enjoy life, and meet the challenges of change through a variety of defined programs and technologies, strong parental and community involvement, and an effective, caring staff."

The British Columbia school system commits itself to develop the individual potential of students. The knowledge, skills and attitudes gained by students are those crucial to a healthy society and a prosperous economy. Responsibility for this goal is shared between the provincial government (through the Ministry of Education) and locally elected school boards.

The Ministry of Education provides direction by identifying three areas of focus for schools. They are:

1. Intellectual Development
2. Social/Human Development
3. Career Development

Attributes of effective public school systems are also identified by the Ministry. The attributes isolated are the following:

1. Accessibility
2. Relevance
3. Equity
4. Quality
5. Accountability*

School Districts operate under these broad guide-lines and develop unique goals intended to meet the needs and concerns of their students and constituents.

We in the Interior School District have stated in clear terms our commitment to the individual needs of your district's young people. Our district is committed to bringing about positive results through processes that the people in the district see as both proper and effective.

* Revised in printing to reflect the Government's new Mandate Statement.

Appendix E3: Brainstorming rules as applied in the Interior School District.**Brainstorming suggestions:**

- ✎ each member should use his or her role (e.g. parent, teacher) when making suggestions.
- ✎ all suggestions should be accepted and written down for later discussion.
- ✎ during this stage of discussion, everyone should have a chance to make suggestions without being rushed or interrupted.
- ✎ once the list is compiled, the potential of each suggested indicator should be discussed:
 - does the indicator fulfill the criteria?
 - how does it fit on the Q-sort?
(Appendix V explains the Q-sort procedure.)

APPENDIX E4

Appendix E4: Sample of Indicator lists from the Interior School District Evaluation.

PROPOSED EFFECTIVENESS CRITERIA

PRODUCT

To value learning:

- o retention rates - night school for dropouts
- o % of students entering post-secondary institutions
e.g. Okanagan College
- o attendance rates
- o involvement in Science Fairs, festivals, etc.

To possess knowledge and skills: (creative independent thinking skills)

- o provincial exam results
- o district exam results

To strive for personal excellence:

- o provincial/A.P./I.B. exam results
- o Math competition - University of Waterloo
- o Passport Program
- o debating program
- o drama/band festivals
- o sports

To experience success:

- o school-based award programs
- o Work Experience Program
- o Board recognition of students
- o Individualized Education Plans students are successful in

APPENDIX E5

Appendix E5: Interior School District Survey Methodology and Design.**1.4. Methodology**

Interior Research used the following process:

- * **Questionnaire Development:** IRC developed a core questionnaire based on consultation with the clients and their consultative committee. Questions of particular significance to each group were added to that group's questionnaire.
- * **Sample Design:** For the general population part of the study, IRC used a random list of telephone numbers to collect a representative sample of 386 respondents.

The sample of 162 employees accurately represents the proportions of employees belonging to the three employee groups, the number of teachers employed at the secondary and elementary levels and the gender split of the staff.

The student sample of 331 accurately represents the proportions of grade eleven and twelve students enrolled in the various schools, as well as the numbers enrolled in grade eleven and twelve and the gender split.

- * **Data Collection:** IRC's trained telephone interviewers collected the data for the employee and general public portions of the study from a supervised central telephone facility. All questionnaires were carefully edited and 10% were verified by a supervisor so that only accurate data is included in the study.

For the student portion of the study the questionnaire was administered in the schools by District staff under the direction of IRC. Budget and time pressures necessitated this less desirable method of data collection.

- * **Data Analysis:** IRC staff entered the data into IRC's computer analysis system and analyzed the data in accordance with the project objectives.

1.5. Level of Confidence

The results of the general population study are accurate to within plus or minus 5%, nineteen times out of twenty.

The results of the student study are accurate to within plus or minus 5%, nineteen times out of twenty.

The results of the staff study are accurate to within plus or minus 7%, nineteen times out of twenty.

APPENDIX F1

Appendix F1: Frequency responses of stakeholder perceptions as to the positive value of the Four-Goal Six-Attribute Framework, and its accuracy.

a. The use of the Four-Goal and Six-Attribute Framework:

positive	1	2	3	4	5	6	7	negative
<u>Inlet</u>	(4)	(3)	(1)	(0)	(0)	(0)	(0)	
<u>South Coast</u>	(2)	(2)	(1)	(0)	(0)	(0)	(0)	
<u>Interior</u>	(4)	(1)	(0)	(0)	(3)	(0)	(0)	
<u>Total</u>	(10)	(6)	(2)	(0)	(3)	(0)	(0)	

b. The use of the Four-Goal and Six-Attribute Framework:

accurate	1	2	3	4	5	6	7	faulty
<u>Inlet</u>	(5)	(1)	(1)	(1)	(0)	(0)	(0)	
<u>South Coast</u>	(1)	(2)	(2)	(0)	(0)	(0)	(0)	
<u>Interior</u>	(4)	(0)	(2)	(1)	(1)	(0)	(0)	
<u>Total</u>	(10)	(3)	(5)	(2)	(1)	(0)	(0)	

APPENDIX F2

Appendix F2: Frequency responses of stakeholders to the development and use of district and ministry key indicators of performance: positive value and accuracy.

a. The development and use of district-based key indicators of educational performance.

positive	1	2	3	4	5	6	7	negative
<u>Inlet</u>	(3)	(4)	(1)	(0)	(0)	(0)	(0)	
<u>South Coast</u>	(2)	(0)	(0)	(3)	(0)	(0)	(0)	
<u>Interior</u>	(6)	(1)	(0)	(0)	(1)	(0)	(0)	
<u>Total</u>	(11)	(5)	(1)	(3)	(1)	(0)	(0)	

b. The development and use of district-based key indicators of educational performance.

accurate	1	2	3	4	5	6	7	faulty
<u>Inlet</u>	(2)	(1)	(4)	(0)	(1)	(0)	(0)	
<u>South Coast</u>	(2)	(0)	(0)	(3)	(0)	(0)	(0)	
<u>Interior</u>	(3)	(2)	(1)	(1)	(1)	(0)	(0)	
<u>Total</u>	(7)	(3)	(5)	(4)	(2)	(0)	(0)	

c. The use of government exam data, retention rates, learning assessment performance of students as provided by the Ministry "key indicators".

positive	1	2	3	4	5	6	7	negative
<u>Inlet</u>	(1)	(2)	(3)	(2)	(0)	(0)	(0)	
<u>South Coast</u>	(3)	(0)	(0)	(2)	(0)	(0)	(0)	
<u>Interior</u>	(5)	(2)	(1)	(0)	(1)	(0)	(0)	
<u>Total</u>	(9)	(4)	(4)	(4)	(0)	(0)	(0)	

Appendix F2 (cont.): Frequency responses of stakeholders to the development and use of key indicators of performance.

d. The use of government exam data, retention rates, learning assessment performance of students as provided by the Ministry's key indicators.

accurate	1	2	3	4	5	6	7	faulty
<u>Inlet</u>	(2)	(1)	(0)	(4)	(1)	(0)	(0)	
<u>South Coast</u>	(3)	(0)	(0)	(2)	(0)	(0)	(0)	
<u>Interior</u>	(4)	(3)	(1)	(0)	(0)	(0)	(0)	
<u>Total</u>	(9)	(4)	(1)	(6)	(1)	(0)	(0)	

e. Frequency of teacher responses compared to non-teacher responses to key indicator qualities.

1. District-based key indicators:

positive	1	2	3	4	5	6	7	negative
<u>Teacher</u>	(0)	(0)	(3)	(2)	(0)	(0)	(0)	
<u>Non-Teacher</u>	(11)	(2)	(0)	(2)	(0)	(0)	(0)	

accurate	1	2	3	4	5	6	7	faulty
<u>Teacher</u>	(0)	(0)	(2)	(2)	(1)	(0)	(0)	
<u>Non-Teacher</u>	(7)	(3)	(3)	(2)	(1)	(0)	(0)	

2. Use of government exam data, retention rates, etc.

positive	1	2	3	4	5	6	7	negative
<u>Teacher</u>	(0)	(0)	(3)	(2)	(0)	(0)	(0)	
<u>Non-Teacher</u>	(9)	(4)	(1)	(2)	(0)	(0)	(0)	

accurate	1	2	3	4	5	6	7	faulty
<u>Teacher</u>	(1)	(0)	(1)	(3)	(0)	(0)	(0)	
<u>Non-Teacher</u>	(8)	(4)	(0)	(3)	(1)	(0)	(0)	

APPENDIX F3

Appendix F3: Frequency responses of stakeholder perceptions with respect to questionnaire desirability and accuracy.

a. The use of surveys of various constituents (parents, non-parents, etc.).

positive	1	2	3	4	5	6	7	negative
<u>Inlet</u>	(4)	(1)	(2)	(0)	(0)	(1)	(0)	
<u>South Coast</u>	(3)	(2)	(0)	(0)	(0)	(0)	(0)	
<u>Interior</u>	(5)	(3)	(0)	(0)	(0)	(0)	(0)	
<u>Total</u>	(13)	(6)	(2)	(0)	(0)	(1)	(0)	

b. The use of surveys of various constituents (parents, non-parents, etc.).

accurate	1	2	3	4	5	6	7	faulty
<u>Inlet</u>	(2)	(1)	(1)	(2)	(1)	(1)	(0)	
<u>South Coast</u>	(0)	(0)	(0)	(3)	(1)	(0)	(1)	
<u>Interior</u>	(4)	(3)	(0)	(1)	(0)	(0)	(0)	
<u>Total</u>	(6)	(4)	(1)	(6)	(2)	(1)	(1)	

APPENDIX F4

Appendix F4: Frequency responses of stakeholder perceptions of evaluation descriptors re efficiency and effectiveness of the evaluation processes and procedures.

Descriptor: Frequency distribution of rating:

1. efficient structure:

agree	1	2	3	4	5	6	7	disagree
<u>Inlet</u>	(4)	(3)	(1)	(0)	(0)	(0)	(0)	
<u>South Coast</u>	(0)	(2)	(1)	(1)	(1)	(0)	(0)	
<u>Interior</u>	(2)	(4)	(1)	(0)	(1)	(0)	(0)	
<u>Total</u>	(6)	(9)	(3)	(1)	(2)	(0)	(0)	

2. functional effectiveness:

agree	1	2	3	4	5	6	7	disagree
<u>Inlet</u>	(3)	(2)	(2)	(0)	(1)	(0)	(0)	
<u>South Coast</u>	(0)	(3)	(2)	(0)	(0)	(0)	(0)	
<u>Interior</u>	(2)	(4)	(0)	(0)	(1)	(1)	(0)	
<u>Total</u>	(5)	(9)	(4)	(0)	(2)	(1)	(0)	

APPENDIX F5

Appendix F5: Frequency responses of stakeholder perceptions of descriptors re information sharing qualities of the evaluation processes and procedures.

<u>Descriptor:</u>	<u>Frequency distribution of rating:</u>							
agree	1	2	3	4	5	6	7	disagree
1. open:								
<u>Inlet</u>	(6)	(0)	(2)	(0)	(0)	(0)	(0)	(0)
<u>South Coast</u>	(2)	(1)	(2)	(0)	(0)	(0)	(0)	(0)
<u>Interior</u>	(5)	(3)	(0)	(0)	(0)	(0)	(0)	(0)
<u>Total</u>	(13)	(4)	(4)	(0)	(0)	(0)	(0)	(0)
2. interesting:								
<u>Inlet</u>	(3)	(5)	(0)	(0)	(0)	(0)	(0)	(0)
<u>South Coast</u>	(2)	(3)	(0)	(0)	(0)	(0)	(0)	(0)
<u>Interior</u>	(3)	(3)	(2)	(0)	(1)	(1)	(0)	(0)
<u>Total</u>	(8)	(11)	(2)	(0)	(0)	(0)	(0)	(0)
3. educative:								
<u>Inlet</u>	(4)	(3)	(1)	(0)	(0)	(0)	(0)	(0)
<u>South Coast</u>	(1)	(3)	(1)	(0)	(0)	(0)	(0)	(0)
<u>Interior</u>	(3)	(4)	(0)	(0)	(1)	(0)	(0)	(0)
<u>Total</u>	(8)	(10)	(2)	(0)	(1)	(0)	(0)	(0)
4. rational:								
<u>Inlet</u>	(5)	(3)	(0)	(0)	(0)	(0)	(0)	(0)
<u>South Coast</u>	(0)	(2)	(2)	(0)	(0)	(0)	(1)	(1)
<u>Interior</u>	(4)	(3)	(0)	(0)	(1)	(0)	(0)	(0)
<u>Total</u>	(9)	(8)	(2)	(0)	(1)	(0)	(1)	(1)

APPENDIX F6

Appendix F6: Frequency distribution of words chosen as three words that best describe stakeholders' role (descriptors of each individual's role (Ind); Team's role (Team)).

Word:	<u>Inlet</u>		<u>South Coast</u>		<u>Interior</u>		<u>Total</u>	
	frequency:		frequency:		frequency:		frequency:	
	Ind	Team	Ind	Team	Ind	Team	Ind	Team
cooperating	(4)	(6)	(3)	(3)	(4)	(2)	(11)	(11)
partnership	(4)	(2)	(2)	(2)	(4)	(4)	(10)	(8)
involved	(2)	(2)	(2)	(2)	(5)	(2)	(9)	(6)
trusted	(4)	(1)	(1)	(0)	(0)	(1)	(5)	(2)
positive	(2)	(5)	(0)	(2)	(2)	(3)	(4)	(10)
active	(1)	(1)	(1)	(1)	(2)	(3)	(4)	(5)
participant	(0)	(2)	(2)	(2)	(1)	(4)	(4)	(3)
important	(1)	(0)	(0)	(0)	(2)	(2)	(3)	(2)
collaborator	(1)	(0)	(2)	(0)	(0)	(0)	(3)	(0)
powerful	(1)	(1)	(0)	(0)	(1)	(0)	(2)	(1)
decisionmaker	(1)	(0)	(0)	(0)	(1)	(1)	(2)	(1)
knowledgable	(0)	(1)	(1)	(2)	(0)	(2)	(1)	(5)
director	(1)	(0)	(0)	(0)	(0)	(1)	(1)	(1)
reticent	(1)	(0)	(0)	(0)	(0)	(0)	(1)	(0)
passive	(1)	(0)	(0)	(0)	(0)	(0)	(1)	(0)
managing	(0)	(0)	(1)	(0)	(0)	(0)	(1)	(0)
shy	(0)	(0)	(0)	(0)	(1)	(0)	(1)	(0)
opinionated	(0)	(0)	(0)	(1)	(0)	(1)	(0)	(2)
compliant	(0)	(0)	(0)	(2)	(0)	(0)	(0)	(2)
strong	(0)	(1)	(0)	(0)	(0)	(0)	(0)	(1)
yielding	(0)	(0)	(0)	(0)	(0)	(1)	(0)	(1)

For the Team choices in the Inlet district, one individual used the term "honest" and another gave only two choices.

APPENDIX F7

Appendix F7: Frequency distribution of word categorization-- describes role well (w), describes role somewhat (s), describes role not at all (n).

Word:	<u>Inlet District</u>			<u>South Coast District</u>			<u>Interior District</u>			<u>Total</u>		
	frequency:			frequency:			frequency:			frequency:		
	w	s	n	w	s	n	w	s	n	w	s	n
cooperating	(8)	(0)	(0)	(5)	(0)	(0)	(8)	(0)	(0)	(21)	(0)	(0)
involved	(7)	(1)	(0)	(5)	(0)	(0)	(8)	(0)	(0)	(20)	(1)	(0)
participant	(7)	(1)	(0)	(5)	(0)	(0)	(8)	(0)	(0)	(20)	(1)	(0)
positive	(8)	(0)	(0)	(4)	(1)	(0)	(6)	(2)	(0)	(18)	(3)	(0)
collaborator	(7)	(1)	(0)	(4)	(1)	(0)	(7)	(1)	(0)	(18)	(3)	(0)
partnership	(6)	(2)	(0)	(5)	(0)	(0)	(7)	(1)	(0)	(18)	(3)	(0)
active	(7)	(1)	(0)	(4)	(1)	(0)	(6)	(2)	(0)	(17)	(4)	(0)
trusted	(6)	(2)	(0)	(5)	(0)	(0)	(5)	(3)	(0)	(16)	(5)	(0)
important	(4)	(3)	(1)	(2)	(2)	(1)	(5)	(3)	(0)	(11)	(8)	(2)
knowledgable	(4)	(4)	(0)	(2)	(2)	(1)	(3)	(5)	(0)	(9)	(11)	(1)
decisionmaker	(3)	(5)	(0)	(2)	(3)	(0)	(3)	(5)	(0)	(8)	(13)	(0)
strong	(4)	(3)	(1)	(1)	(3)	(1)	(3)	(5)	(0)	(8)	(11)	(2)
managing	(2)	(4)	(2)	(3)	(2)	(0)	(3)	(3)	(2)	(8)	(9)	(4)
director	(1)	(3)	(4)	(1)	(2)	(2)	(3)	(2)	(3)	(5)	(7)	(9)
observer	(2)	(5)	(1)	(1)	(3)	(0)	(0)	(5)	(3)	(3)	(13)	(4)
powerful	(2)	(2)	(4)	(0)	(4)	(1)	(1)	(4)	(3)	(3)	(10)	(8)
opinionated	(0)	(2)	(6)	(1)	(1)	(3)	(2)	(6)	(0)	(3)	(9)	(9)
controller	(1)	(4)	(3)	(1)	(4)	(0)	(0)	(4)	(4)	(2)	(12)	(7)
yielding	(0)	(4)	(0)	(1)	(3)	(1)	(1)	(5)	(2)	(2)	(12)	(7)
indulged	(0)	(1)	(7)	(0)	(3)	(2)	(2)	(4)	(2)	(2)	(8)	(11)
shy	(1)	(1)	(6)	(0)	(1)	(4)	(1)	(0)	(7)	(2)	(2)	(17)
compliant	(0)	(4)	(4)	(0)	(5)	(0)	(1)	(7)	(0)	(1)	(16)	(4)
follower	(0)	(3)	(5)	(0)	(4)	(1)	(1)	(5)	(2)	(1)	(12)	(8)
acquiescing	(1)	(2)	(5)	(0)	(4)	(1)	(0)	(3)	(5)	(1)	(9)	(11)
subordinate	(0)	(2)	(6)	(1)	(3)	(1)	(0)	(4)	(4)	(1)	(9)	(11)
reticent	(1)	(0)	(7)	(0)	(2)	(3)	(0)	(3)	(5)	(1)	(5)	(15)
passive	(1)	(1)	(6)	(0)	(1)	(4)	(0)	(1)	(7)	(1)	(3)	(17)
insignificant	(1)	(0)	(7)	(0)	(0)	(5)	(0)	(0)	(8)	(1)	(0)	(20)
appeasing	(0)	(2)	(6)	(0)	(5)	(0)	(0)	(4)	(4)	(0)	(11)	(10)
ignorant	(0)	(2)	(6)	(0)	(2)	(3)	(0)	(4)	(4)	(0)	(8)	(13)
powerless	(0)	(0)	(8)	(0)	(2)	(3)	(0)	(2)	(6)	(0)	(4)	(17)
negative	(0)	(1)	(7)	(0)	(1)	(4)	(0)	(0)	(8)	(0)	(2)	(19)
threatened	(0)	(0)	(8)	(0)	(2)	(3)	(0)	(0)	(8)	(0)	(2)	(19)
weak	(0)	(0)	(8)	(0)	(1)	(4)	(0)	(0)	(8)	(0)	(1)	(20)
inactive	(0)	(0)	(8)	(0)	(0)	(5)	(0)	(0)	(8)	(0)	(0)	(21)

APPENDIX F8

Appendix F8: Frequency distribution of stakeholder perceptions of evaluation descriptors re ethical quality.

<u>Descriptor:</u>	<u>Frequency Distribution of Rating:</u>							
agree	1	2	3	4	5	6	7	disagree
1. just:								
<u>Inlet</u>	(5)	(1)	(2)	(0)	(0)	(0)	(0)	
<u>South Coast</u>	(1)	(0)	(1)	(2)	(1)	(0)	(0)	
<u>Interior</u>	(5)	(3)	(0)	(0)	(0)	(0)	(0)	
<u>Total</u>	(11)	(4)	(3)	(2)	(1)	(0)	(0)	
2. open:								
<u>Inlet</u>	(6)	(0)	(2)	(0)	(0)	(0)	(0)	
<u>South Coast</u>	(2)	(1)	(2)	(0)	(0)	(0)	(0)	
<u>Interior</u>	(5)	(3)	(0)	(0)	(0)	(0)	(0)	
<u>Total</u>	(13)	(4)	(4)	(0)	(0)	(0)	(0)	
3. honest:								
<u>Inlet</u>	(4)	(2)	(2)	(0)	(0)	(0)	(0)	
<u>South Coast</u>	(1)	(3)	(1)	(0)	(0)	(0)	(0)	
<u>Interior</u>	(4)	(3)	(0)	(0)	(1)	(0)	(0)	
<u>Total</u>	(9)	(8)	(3)	(0)	(1)	(0)	(0)	
4. has integrity:								
<u>Inlet</u>	(4)	(1)	(3)	(0)	(0)	(0)	(0)	
<u>South Coast</u>	(3)	(0)	(2)	(0)	(0)	(0)	(0)	
<u>Interior</u>	(3)	(4)	(0)	(0)	(1)	(0)	(0)	
<u>Total</u>	(10)	(5)	(5)	(0)	(1)	(0)	(1)	

APPENDIX F9

Appendix F9: Frequency distribution of stakeholder perceptions of evaluation descriptors re authority relationships.

<u>Descriptor:</u>	<u>Frequency Distribution of Rating:</u>							
	1	2	3	4	5	6	7	disagree
1. dignifying of constituent opinions:								
<u>Inlet</u>	(5)	(1)	(2)	(0)	(0)	(0)	(0)	
<u>South Coast</u>	(3)	(0)	(2)	(0)	(0)	(0)	(0)	
<u>Interior</u>	(4)	(4)	(0)	(0)	(0)	(0)	(0)	
<u>Total</u>	(12)	(5)	(4)	(0)	(0)	(0)	(0)	
2. minimal use of executive authority:								
<u>Inlet</u>	(4)	(0)	(2)	(1)	(1)	(0)	(0)	
<u>South Coast</u>	(2)	(0)	(1)	(1)	(0)	(1)	(0)	
<u>Interior</u>	(3)	(2)	(0)	(1)	(2)	(0)	(0)	
<u>Total</u>	(9)	(2)	(3)	(3)	(3)	(1)	(0)	
3. politics conducted responsibly:								
<u>Inlet</u>	(2)	(4)	(1)	(1)	(0)	(0)	(0)	
<u>South Coast</u>	(1)	(3)	(1)	(0)	(0)	(0)	(0)	
<u>Interior</u>	(4)	(3)	(1)	(0)	(0)	(0)	(0)	
<u>Total</u>	(7)	(10)	(3)	(1)	(0)	(0)	(0)	
4. freedom within guidelines:								
<u>Inlet</u>	(5)	(2)	(1)	(0)	(0)	(0)	(0)	
<u>South Coast</u>	(2)	(0)	(1)	(0)	(2)	(0)	(0)	
<u>Interior</u>	(4)	(3)	(1)	(0)	(0)	(0)	(0)	
<u>Total</u>	(11)	(5)	(3)	(0)	(2)	(0)	(0)	

APPENDIX F10

Appendix F10: Frequency responses of Stakeholder Team members to possible achievements of the evaluation process.

<u>Descriptor:</u>	<u>Frequency Distribution of Rating:</u>							
agree	1	2	3	4	5	6	7	disagree
1. provides a springboard for effective and long term planning:								
<u>Inlet</u>	(4)	(2)	(1)	(1)	(0)	(0)	(0)	
<u>South Coast</u>	(3)	(0)	(0)	(1)	(0)	(1)	(0)	
<u>Interior</u>	(3)	(3)	(1)	(0)	(1)	(0)	(0)	
<u>Total</u>	(10)	(5)	(2)	(2)	(1)	(1)	(0)	
2. provides a consistent knowledge base for members of the management team to make better decisions:								
<u>Inlet</u>	(5)	(0)	(3)	(0)	(0)	(0)	(0)	
<u>South Coast</u>	(1)	(0)	(2)	(1)	(0)	(1)	(0)	
<u>Interior</u>	(4)	(1)	(2)	(0)	(1)	(0)	(0)	
<u>Total</u>	(10)	(1)	(7)	(1)	(1)	(1)	(0)	
3. helps to provide a clear picture, or vision, of where the education system is and should go in the future:								
<u>Inlet</u>	(2)	(2)	(3)	(0)	(1)	(0)	(0)	
<u>South Coast</u>	(0)	(1)	(1)	(1)	(0)	(1)	(1)	
<u>Interior</u>	(2)	(4)	(0)	(1)	(1)	(0)	(0)	
<u>Total</u>	(4)	(7)	(4)	(2)	(2)	(1)	(1)	

APPENDIX F11

Appendix F11: Frequency responses of Stakeholder Team members to possible achievements of the evaluation process.

<u>Descriptor:</u>	<u>Frequency Distribution of Rating:</u>							
agree	1	2	3	4	5	6	7	disagree

1. assesses student performance, or achievement, rather than school processes and procedures:

<u>Inlet</u>	(0)	(4)	(1)	(0)	(2)	(1)	(0)
<u>South Coast</u>	(1)	(1)	(2)	(0)	(1)	(0)	(0)
<u>Interior</u>	(3)	(2)	(1)	(1)	(1)	(0)	(0)
<u>Total</u>	(4)	(7)	(4)	(1)	(4)	(0)	(0)

2. will improve learning for students:

<u>Inlet</u>	(0)	(1)	(4)	(2)	(0)	(1)	(0)
<u>South Coast</u>	(2)	(0)	(1)	(1)	(0)	(0)	(1)
<u>Interior</u>	(0)	(3)	(2)	(2)	(1)	(0)	(0)
<u>Total</u>	(2)	(4)	(7)	(5)	(1)	(1)	(1)

APPENDIX F12

Appendix F12: Frequency responses of Stakeholder Team perceptions of impact of the evaluation in terms of gaining public support for education:

<u>Descriptor:</u>	<u>Frequency Distribution of Rating:</u>							
agree	1	2	3	4	5	6	7	disagree
1. The evaluation will: develop a higher degree of shared commitment and purpose amongst parents and the general public:								
<u>Inlet</u>	(0)	(1)	(3)	(2)	(2)	(0)	(0)	
<u>South Coast</u>	(0)	(1)	(3)	(0)	(0)	(1)	(0)	
<u>Interior</u>	(0)	(3)	(5)	(0)	(0)	(0)	(0)	
<u>Total</u>	(0)	(5)	(11)	(2)	(2)	(1)	(0)	

APPENDIX F13

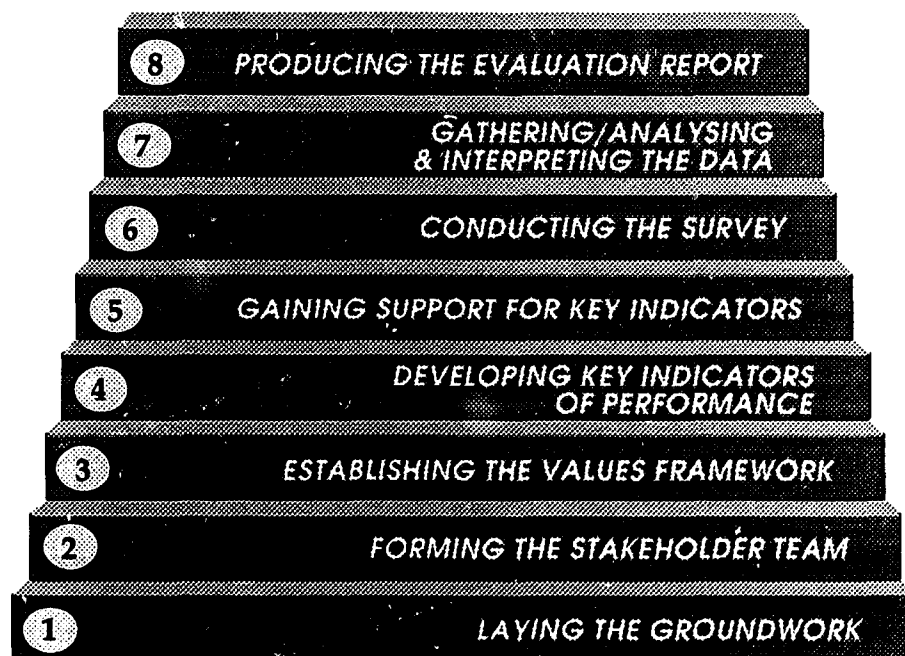
Appendix F13: Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

CONDUCTING A DISTRICT EVALUATION

This section describes the steps in conducting an evaluation. The procedure for carrying out each step is provided, followed by suggestions, drawn from the pilot studies, on how to carry out each step most efficiently.

This section also includes descriptions of how to use the evaluation information to gain support and feedback from stakeholder groups and to plan district improvements.

The Eight Steps of District Evaluation:



Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

1

LAYING THE GROUNDWORK

Good preparation for the evaluation will make each of the following steps easier to complete. This involves providing information about the evaluation, gaining a commitment from everyone involved, planning each event and establishing how the evaluation information will be used.

Procedure for Step 1

- The school board and district management team make a commitment to conduct the evaluation and to provide the time and resources it needs.
- The idea is presented to all major interest groups: teachers, administrators, the school board, district staff, parent groups and other interest groups. Each group is asked to notify its constituents and to send a representative to a preliminary stakeholder meeting.
- Information about the evaluation and its purposes is presented to the public.
- At the preliminary stakeholder meeting, the evaluation and its uses are outlined so that they are clearly understood by all stakeholder groups.
- A time line for events is established.
- An agreement is reached on how the evaluation results will be used to serve district needs.
- Follow-up plans are made for keeping all stakeholder groups informed after the evaluation.

Suggestions for Step 1

- ✓ Make sure that all groups who will be involved clearly understand the purpose and process of the evaluation.
- ✓ Choose a time frame for the evaluation that will fit well with the activities of those involved.
- ✓ The more information people receive about the evaluation, the more likely they are to understand its purpose and to give it their support.
- ✓ Allow lots of time for preparation.

Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

2

FORMING THE STAKEHOLDER TEAM

The evaluation is conducted by a team made up of representatives from all stakeholder groups. These include the school board, teachers' union, custodial and clerical union, students parent groups, administrator's association, district staff organizations and public interest groups. The stakeholder collaborative approach ensures a commitment from all interest groups and a teamwork approach to decision-making.

Procedure for Step 2

- At the preliminary stakeholder meeting, each group is asked to select a representative to join the team.

Team duties:

- *creating the values framework*
- *identifying key indicators*
- *collecting/analyzing interpreting data*
- *developing the survey*
- *producing the report*

- A meeting of the stakeholder team is held to explain the duties of the team.
- Time lines for each event, suitable to all team members, are agreed on.
- Specific job responsibilities are assigned.

Suggestions for Step 2

- ✓ Stakeholder team members should be in positions within their constituent groups that will allow them to receive input and to keep their members informed about the evaluation.
- ✓ Be sure that team members understand the time and energy commitment required.
- ✓ Make sure that each team member has opportunities to be heard and to take part in decision-making. Structuring small group sessions in which leadership can rotate helps to ensure that all points of view are considered.
- ✓ To encourage commitment, cooperation and collective ownership of the process, no one should be in a position of complete authority over the team.

Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

3

ESTABLISHING THE VALUES FRAMEWORK

The framework chosen represents the district's goals and ideals. These form the beginning point for analyzing the district's educational performance.

Procedure for Step 3

- The team decides whether to use the Ministry's 3 goals and 5 attributes framework or an existing mission statement. (Appendix I provides a complete description of the Ministry framework.)
- If an existing mission statement is used, the team examines it to make sure that it covers all areas of the Ministry framework and makes any necessary adjustments.
- The team may choose to adapt the language of the Ministry framework to express the priorities and individuality of its community.

GOALS AND ATTRIBUTES OF EDUCATION

Goals of Education	Intellectual Development Human/Social Development Career Development
Attributes of Education	Accessibility Quality Equity Relevance Accountability

Suggestions for Step 3

- ✓ Before beginning, the team should have a thorough understanding of the Government's goals and attributes framework.

Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

4

DEVELOPING KEY INDICATORS OF PERFORMANCE

Key indicators of performance measure the achievement of the district's ideals. The stakeholder team decides which indicators most accurately express the state of achievement for each of its ideals. Both qualitative (e.g. existing data) and quantitative (e.g. survey results) indicators are developed for each element of the district's values framework. Together they provide a clear picture of the district's strengths and weaknesses and provide a base for monitoring progress.

Procedure for Step 4

- The team members are informed about the nature of key indicators and the criteria for their development. (Appendix II lists the criteria for key indicator selection.)
- Either before or during this stage of the evaluation, each team member meets with his or her constituents to gather input on the choice of key indicators (see Step 5).
- The team brainstorms on which indicators are suitable for each element of the values framework.
- The team decides upon both qualitative and quantitative indicators for each element. (Appendix III provides a list of possible key indicators.)
- The team determines how data will be collected and presented.
- The responsibility for collecting data for each indicator is assigned, appropriate to members' roles.
- Potential data sources are identified and refinements to key indicator development is made based on available data. (Appendix IV provides some possible data sources.)
- A date is set for having data ready.

Suggestions for Step 4

- ✓ Before starting, be sure that team members thoroughly understand the concept of key indicators.
- ✓ Choose indicators that can be clearly defined and represented.

Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

Suggestions for Step 4 (cont'd)

✓ This stage of the evaluation is time-consuming. Make sure the team has enough time to choose indicators that meet all the criteria.

✓ Try to ensure that indicators cover the achievements of all students.

✓ Consider the practicality of finding appropriate data for each indicator. Indicators that cannot be expressed in measurable terms may be represented by 'demonstration' indicators, e.g. photographs, lists of scholarship achievements, video yearbooks.

Brainstorming suggestions:

- ✎ each member should use his or her role (e.g. parent, teacher) when making suggestions.
- ✎ all suggestions should be accepted and written down for later discussion.
- ✎ during this stage of discussion, everyone should have a chance to make suggestions without being rushed or interrupted.
- ✎ once the list is compiled, the potential of each suggested indicator should be discussed:
 - does the indicator fulfill the criteria?
 - how does it fit on the Q-sort?
(Appendix V explains the Q-sort procedure.)

Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

5

GAINING SUPPORT FOR KEY INDICATORS

The input of stakeholder constituent groups is important in choosing key indicators that represent the widest possible range of values and viewpoints. Giving people the chance to contribute their intuitions about how educational success is measured lays the basis for empowerment and future involvement.

Team members are responsible for educating their groups about the purpose and nature of key indicators, gathering input from them and presenting it for inclusion in the team's decisions.

Procedure for Step 5

- The team decides either to meet with constituent groups before the team's own brainstorming session or to take the outcome of that session to their groups for discussion.
- If the former option is chosen, team members meet with their groups to explain key indicator development and to brainstorm for suggestions using the criteria and considerations outlined in Step 4. The indicators selected are brought back to the team for discussion in its choice of indicators.
- If the latter option is chosen, members meet with their groups to present the team's selection for discussion. Suggestions and differences of opinion should be reported back to the team for consideration during the refinement of the key indicator selection.
- In either case, constituent groups should remain involved during this stage, receiving feedback on their key indicator selections and having the opportunity to give further input until a consensus is reached.

Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

Suggestions for Step 5

- ✓ Meetings of this type need good management. Useful strategies include: setting time limits on discussions, dividing the work among sub-groups that report back to the main group, and using a Q-sort in the selection process. (See Appendix V for the Q-sort procedure.)
- ✓ Try to ensure that all viewpoints are heard, that people feel their input is valued, and that the group receives feedback on their suggestions.

Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

6

CONDUCTING THE SURVEY

An opinion survey contributes subjective data to the evaluation, providing a more complete picture of the district's performance. The stakeholder team decides what information they need and from whom. The survey should be conducted by a professional polling company.

procedure for step 6

- The team determines the goals of the survey, who will be surveyed, and the time line for carrying out the survey.
- A set of survey questions is developed. These should be comparable across constituencies and relevant to the goals of the evaluation.
- A professional researcher is hired to produce the questionnaire, conduct the survey and collect the results.
- The questionnaire is submitted to the team for adjustments and approval before being conducted.

Suggestions for Step 6

- ✓ Allow adequate time for the team to decide who should be surveyed and what information should be gathered.
- ✓ Ideally a survey should be conducted every 3 years. Survey questions should be designed to be comparable so that trends can be identified.

Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

7

GATHERING/ANALYZING/INTERPRETING THE DATA

During this stage data from all sources are collected and interpreted for presentation. The team checks that the data are both subjective and objective, that there are adequate data for each indicator, and that data are accurate and relevant to the values framework.

Procedure for Step 7

- The team meets on a prearranged date to submit all data collected. These include quantitative data, survey results and characteristics of the district that add to interpretation of the data
- Data are sorted according to each framework element and organized into four categories:
 - data indicating a significant strength
 - data indicating a significant weakness
 - data that show significant differences in the responses of different groups
 - data that add to an understanding of the district.
 (Appendix VI provides a detailed description of this procedure.)
- Data that do not fit in to these categories and that do not add unique, pertinent information about the district are discarded.
- The team checks that:
 - all framework elements are accurately represented
 - both subjective and objective data are available
 - there are enough key indicators for each framework element.
- If any framework element is found to be insufficiently covered, a team member is assigned to generate other key indicators.

Suggestions for Step 7

- ✓ Organizing, analyzing and interpreting data takes time. The team should arrange to set aside enough time to deal with this part of the evaluation thoroughly.
- ✓ Accuracy of the data is crucial to the evaluation. Data should never be tampered with and data that reflect weakness in the district should never be disregarded because of this fact.

Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

8

PRODUCING THE EVALUATION REPORT

The evaluation report presents key indicators of performance to express the level of achievement of the district's values framework. The report is the final stage in the evaluation and the first stage in using the information for district improvement.

The actual production of the report is left up to the district. The district also decides who should receive the report and whether different versions of the report should be produced for different audiences.

Procedure for Step 8

- The team decides on intended audiences, length, format and purpose of the report(s) and makes plans for its distribution.
- Responsibilities for writing, graphics, layout and printing are assigned.
- The team examines drafts of the report for readability, integrity and presentation of the information and approves it for printing.

Suggestions for Step 8

- ✓ The style of presentation and level of readability of the report should be appropriate to its intended audience. A report intended for public information purposes may not be suitable for use by district administrators.
- ✓ An attractive, readable, well layed-out report encourages people to read it and thereby to gain understanding about the state of education in the district.
- ✓ The following sections should be included in the report:
 - a statement of purpose/vision
 - characteristics of the district
 - definitions of key elements, e.g., key indicators.

Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

USING THE RESULTS OF FORMATIVE EVALUATION

An evaluation report is often regarded as the final step in a program or plan. In a formative evaluation, the report is seen as the first step in the process of improving district performance.

How the evaluation information will be used, by whom and for what purposes should be decided before the evaluation takes place. A clearly defined follow-up plan formalizes the district's commitment to use the information and helps to ensure that time and resources needed for follow-up are available.

Two stages of follow-up activity are suggested:

1. Using the information to gain support and input from stakeholder groups and the public.
2. Using this input and the evaluation information as a basis for planning.

Follow-up: Distributing the report and gathering input

Each district makes its own plans for presenting the evaluation information. The following is a list of activities that were carried out by school districts taking part in the pilot studies.

School board

- A school board meeting was held to analyze the report. The board went through the report systematically, identifying areas of concern in order to design an action plan for improvement.
- A school board committee was organized to develop a strategy for improving relations between the district and its non-parent and employer constituents.

Schools

- A meeting between the superintendent and principals was held to discuss how to communicate the information to parent groups and staff.

Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

Parent groups

- A meeting of the parent advisory committee was held. The outcome of the discussion was summarized and submitted to the school board.
- A meeting of each school's parent group was held to discuss the report.

Students

- Reports were sent to the district's secondary English departments. Students' written papers on the state of education in the district were later submitted to the superintendent.

Employers

- Copies of the report were sent to all employers. A meeting was held to talk about their concerns.

Public

- A video was made of the report highlights. The video was shown at the local mall during education week, clips of it were shown on local t.v, and it plays in the foyer of the district office.
- A newspaper article was written about the evaluation.
- A newspaper insert summarizing the report was included in the local paper.
- Reports were placed in dentists' and doctors' offices and distributed to real estate companies for individuals moving into the district.
- Stakeholder team members presented the report and video to local service clubs.

Suggestions for presenting the evaluation information

- 📎 Give the evaluation follow-up a high public profile. Media releases, meetings and other activities should be well-communicated to everyone concerned.
- 📎 Make sure that all interest groups are given ample opportunity through a structured input process to provide feedback and suggestions for improvement.
- 📎 Possibly the role of the stakeholder team should be expanded to include the responsibility for communicating the information to all important groups.

Appendix F13 (cont.): Outline of the Information Profile System as presented to School Districts at the conclusion of the pilot study.

Follow-up: Planning

Without a clearly defined series of activities for converting evaluation information into long range plans, the evaluation's potential goals are only partially achieved. An evaluation itself can provide direction and focus, but if these are not constrained through a commitment to organized decision-making, the impetus for long term planning is soon lost.

How the evaluation information will be used in planning improvements is left up to the district. The following is a possible framework for organizing the planning stages.

1. *Review current practices:*
 - conduct formative evaluation
 - provide information to major stakeholders
 - receive input from major stakeholders.

2. *Define the preferred future:*
 - consider reaction from major stakeholders to current practices
 - establish the district's major educational goals
 - prioritize the goals

3. *Set a plan for reaching the preferred future:*
 - plan the budget
 - establish short and long term time lines
 - develop strategies for implementation
 - specify roles of all major stakeholders

APPENDIX F14

Appendix F14: The B.C. Government's Mandate Statement for Education.

Part C: Policy Statement on Public Schools

Goals of Education

Prime Goal of Public Schools -- Supported by the Family & Community

• *Intellectual Development* -- to develop the ability of students to analyze critically, reason and think independently, and acquire basic learning skills and bodies of knowledge; to develop in students a lifelong appreciation of learning, a curiosity about the world around them and a capacity for creative thought and expression.

Goals that are shared among Schools, the Family & Community

Schools are expected to play a major role, through learning experiences and supervised practice, in helping students to achieve the following goals:

• *Human & Social Development* -- to develop in students a sense of self-worth and personal initiative; to develop an appreciation of the fine arts and an understanding of cultural heritage; to develop an understanding of the importance of physical health and well being; to develop a sense of social responsibility, and a tolerance and respect for the ideas and beliefs of others.

• *Career Development* -- to prepare students to attain their career and occupational objectives; to assist in the development of effective work habits and the flexibility to deal with change in the workplace.

Attributes of the Public School System:

• *Accessibility* -- a variety of programs is available in the province to meet the full range of student needs.

• *Relevance* -- programs are current, and relevant to the needs of the learner.

Appendix F14 (cont.): The B.C. Government's Mandate Statement for Education.

- *Equity* -- resources are allocated fairly.
- *Quality* -- professional teaching and administration are of high quality.
- *Accountability* -- resources are allocated in a cost-effective manner; parents and the community are informed of the progress of schools and are involved as partners in planning.

Duties, Rights and Responsibilities

Students: have the opportunity to avail themselves of a quality education consistent with their abilities, the opportunity to share in the shaping of their educational programs, and the opportunity to determine their career and occupational goals. They have a responsibility to make the most of their opportunities, to respect the rights of others, and to cooperate with fellow students in the achievement of their goals.

Parents: have the right and responsibility to participate in the process of determining the educational goals, policies and services provided for their children. They have a primary responsibility to ensure that children are provided with the healthy and supportive environment necessary for learning. They have a responsibility to help shape and support the goals of the school system and to share in the tasks of educating their young.

Teachers: have the right to exercise professional judgment in providing instruction to students in accordance with specified duties and powers. They have a corresponding responsibility to ensure that each student is provided with quality instruction, to participate in all normal school activities and to monitor the behavior and progress of each learner in accordance with provincial and local policies. They have a responsibility to communicate with students and parents, and are accountable to the School Board and its delegates.

School Principals: have the right to exercise professional judgment in managing the school in accordance with specified duties and powers. They have a corresponding responsibility to ensure that each student is provided with opportunities for a quality education. Principals are to provide administrative leadership, in consultation with teachers and the community, that reflects the aspirations of parents and the school community and that is consistent with provincial and district guidelines. They cooperate with parents and the community in the delivery of non-educational support services to students, and focus on the following areas of school concern: (1) student access and achievement; (2) quality teaching; (3) communication with parents and the community; and (4) accountability to parents and to the Board.

School Boards: have a duty to govern districts and their schools in accordance with

APPENDIX G1

Appendix G.1: Steps of Value judgement experienced in the first stage of the Inlet evaluation.

Stage 1: Measurement

Level 1: Developing measures of performance

1. choice of framework;
2. choice of def'n of framework elements, title and terms.
3. choice of key indicator definition.
4. brainstorming of general key indicator concepts.
5. endorsement of key indicator suggestions.
6. application of key indicator definition to key indicators; judgment as to acceptability.
7. culling of key indicators.
8. defining key indicators in a measurable way.
9. choice of questionnaire questions
10. questionnaire wording.
11. choice of constituent groups to question.
12. choice of sample/method.

APPENDIX G2

Appendix G2: Employing Hodgkinson's (1978, 1983) Value Paradigm to characterize value priorities in the formative evaluation model:

Level I:	Improvement	(ideological value)
Level II:		
	A. Values of democracy and education embodied in Goals and Attributes of Education System.	(moral values)
	B. Key indicators of performance designed to measure goal and attribute performance; criteria for key indicators and research information	(consensual values)
Level III:	Stakeholder values, collected through questionnaires; key indicator process.	(preferential values)

APPENDIX G3

Appendix G3: Q-sort ratings of stakeholder opinions of the main factors accounting for the success of the district evaluation--in order of priority.

1. Senior Administrators treated committee members with respect and consideration.

Q-sort that allowed equal input in decision making.

The use of "judgement and evaluation" at each stage of decision making throughout the process.

2. Time lines and actions that were well-structured

3. Clearly defined tasks.

Reliance on knowledge and values of interest groups.

4. Alterations of conventional stakeholder roles

5. Novelty of process.

6. Senior administrators knowing what they wanted and persuading people to do what they wanted.