

EDUCATIONAL EVALUATION: TWO THEORETICAL MODELS

IN A CORPORATE BASED APPLICATION

By

GORDON W. BARRETT

B.Ed., The University of British Columbia

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(Department of Language Education)

We accept this thesis as conforming

to the ~~the~~ required standard

THE UNIVERSITY OF BRITISH COLUMBIA

APRIL 1998

© GORDON WILLIAM BARRETT, 1998

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

Department of LANGUAGE EDUCATION

The University of British Columbia
Vancouver, Canada

Date 25 APRIL 1998

Abstract

The Provus Discrepancy Evaluation Model (1973) and the Stufflebeam et al. C.I.P.P. Evaluation Model (1973) are examined against the backdrop of two evaluations that were conducted by unskilled evaluators in a corporate based setting by a large corporation. Differences between the two theoretical models and the two corporate evaluations revealed that there are factors, not considered in the theoretical models, which can impact their effectiveness when practically applied. The Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model were in some ways appropriate and both committees would have benefited from utilizing similar evaluation models. Failure of these two committees to address significant aspects of formal evaluation might have been remedied by the application of formal evaluation models. Educational evaluation models also have significant "gaps" including personal investment of committee members, corporate agendas, cost and financial impacts, bias and stall points (the point where the evaluation model ceases to be effective in the current context). Corporate evaluations are an ongoing process and require different types of evaluation models, depending upon current need (Stall Point Theory).

Two evaluations that were conducted by task force committees in a corporate setting were examined. One task force committee examined how corporate training was being carried out and the second task force committee examined adherence to how corporate policy was being carried out.

Using a formal evaluation model would have provided structure, objective clarification, and greater confidence in the results and recommendations of the corporate evaluations. The development of educational evaluation models would be enhanced by considering the needs of the end users of the models, making the models more dynamic by increasing flexibility for general/specific application and acknowledging the models' limitations.

TABLE OF CONTENTS

	iv
Abstract	ii
Table of Contents	iv
List of Tables	ix
List of Figures	x
Acknowledgements	xi
Dedication	xii
CHAPTER ONE: INTRODUCTION	1
Statement of the Problem	1
Theoretical Models Examined	2
Discrepancy Evaluation Model	2
C.I.P.P. Evaluation Model	3
Origins of the Corporate Evaluation Committees .	4
Claims Training Evaluation Task Force	
Committee	5
Litigation Management Evaluation Task Force	
Committee	6
Pre-Evaluations Considerations	6
Internal Evaluation Advantages	7
Internal Evaluation Disadvantages	8
External Evaluation Advantages	9
External Evaluation Disadvantages	10
A Combination Approach	11
Significance of the Study	12

	v
Significance to Education	12
Significance to Business	13
Two Questions for Consideration	13
Limitations of This Study	14
Summary	17
CHAPTER TWO: LITERATURE ON EDUCATIONAL EVALUATION ..	19
History of Evaluation in Education	19
Education Evaluation Models Used in This Study .	26
Provus Discrepancy Evaluation Model	26
Advantages	26
Limitations	27
Stufflebeam et al C.I.P.P. Evaluation Model	28
Advantages	28
Limitations	30
Other Education Evaluation Models Considered ...	31
Tylerian Evaluation Approach	31
Hammond's Evaluation Approach	32
U.C.L.A. Evaluation Model	33
Stake's Countenance Evaluation Model	34
Rationale for Selecting the Provus and	
Stufflebeam et al. Evaluation Models	35
Education Evaluation Research Sources	36
Results of Documentation Search	37
Summary	37

	vi
CHAPTER THREE: METHOD	38
Task Force Members	38
Role of the Participant Observer	39
Participants in the Claims Training Evaluation	
Task Force Committee Investigation	40
Criteria for Focus Group Selection	41
Task Force Procedures	42
Methodology: Claims Training Evaluation Task	
Force Committee - Historical Background	43
Methodology: Claims Training Evaluation Task	
Force Committee - Current Situation	44
Litigation Management Evaluation Task Force	
Committee	48
Data Collection	55
Summary	58
CHAPTER FOUR: RESULTS	59
Evaluation Model Application Comparison	60
Stage I - Programme Design	60
Stage II - Programme Operation	66
Stage III - Programme Interim Products	70
Stage IV - Programme Terminal Products	74
Stage V - Programme Cost	79
Stufflebeam et al (1977) C.I.P.P. Evaluation	
Model	81

	vii
Context Evaluation	81
Input Evaluation	82
Process Evaluation	88
Product Evaluation	92
Summary	94
 CHAPTER FIVE: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	96
Summary	97
Identification of Programme Objectives and Placing Them In Context	97
Examination of Programme Structure and Supporting Infrastructure	99
Anticipation of Barriers to Success	101
Orientation of Programme Towards Results ..	102
Conclusions	102
Benefits of Using Provus and Stufflebeam's Models	102
Financial Comparison for Cost Containment .	106
Committees' Efforts Facilitated by Using a Formal Evaluation Model	107
Developing Foundations of Evaluations	115
Evaluation Approaches Must Be Flexible in Practice	122
Stall Points	123

	viii
Recommendations	125
Epilogue	129
Impact of External Influences on Evaluation	
Model Effectiveness	129
External Influences Acting upon the Claims	
Training Evaluation Task Force Committee ..	129
External Influences Acting upon the Litigation	
Management Evaluation Task Force Committee	133
References	138

List of Tables

Table

1.	U.C.L.A. and C.I.P.P. Evaluation Models - Comparative Process	34
2.	Comparison - Provus Discrepancy Evaluation Model, Stage I, and the Litigation Management Evaluation Task Force Committee Mandate	65
3.	Comparison - Provus Discrepancy Evaluation Model, Stage II, and the Litigation Management Evaluation Task Force Committee Mandate	69
4.	Bulletin CDB894 Failure: Reasons For Failure	91

List of Figures

Figure

1. Litigation Management Evaluation Task
Committee Weighting of Evaluation Process..... 118
2. Claims Evaluation Task Force Committee
Weighting of Evaluation Process..... 120

ACKNOWLEDGEMENTS

I would like to thank my wife, Lynda, for all of the help and support that she has given me throughout my studies. She kept me going when I was frustrated and overwhelmed at times. She has helped our family by taking on more than her fair share of the load while helping me to balance work and University. I would also like to thank my children, Rachel and William, for giving me their understanding when I was away at classes for a good part of their lives. My Grandmother, Lilian Barrett started me off on my graduate work many years ago and while she is not alive today I am certain that she is watching as this now comes to a close.

Finally, I would like to thank Dr. Joe Belanger for his patience and support while we tried to reshape my business mind back into an educational focus on a regular basis.

I am writing this acknowledgement on Good Friday. It is a reminder that with every end there is always a new beginning. The promise of a glorious tomorrow with every new moment has been given to us if we are prepared to stand steadfast in faith.

Pax

Gordon William Barrett

April 10, 1998

DEDICATION

I would like to dedicate this thesis to my wife, Lynda and to my parents, Gordon and Georgia. I wrote the Acknowledgement on Good Friday, little knowing that my mother would pass away on Easter Day; the day of the Resurrection. In the Acknowledgement I spoke of the beginnings and endings that we encounter on a daily basis. It is written that God is the Alpha and the Omega, the beginning and the end. My mother's earthly life has come to an end and now she begins a new journey through Paradise. She was always proud of her family and we were proud that she was our mother.

Gordon Barrett

April 20, 1998

CHAPTER ONE

INTRODUCTION

The question this study addresses is what strengths and weaknesses are revealed in two corporate evaluations when they are judged post-hoc using two evaluation models designed for educational evaluation, the Provus Discrepancy Evaluation Model and the Stufflebeam et al. Context, Input, Process, Product Model. A related question is what kinds of revisions in the educational evaluation models do applications in a corporate setting suggest?

This study examines these two theoretical educational evaluation models and compares them to two evaluation models that were, apparently, intuitively developed by evaluation committees for evaluation programmes within a large corporation. An examination of the two theoretical evaluation models and the two intuitively developed evaluation models identifies similarities and differences which lead to a better understanding of the evaluation process that functioned within the context of this study.

Statement of the Problem

This study examines the processes used by the two internal evaluation committees in order to explore the differences and similarities between evaluation models that are developed intuitively by untrained evaluators and

theoretical educational evaluation models. Such a comparison should provide insights for both business and education.

Theoretical Models Examined

Using Worthen and Sanders (1987) text, "Educational Evaluation: Alternative Approaches and Practical Guidelines", and Miller and Seller's (1990) "Curriculum Perspectives and Practices" as a guide, several theoretical evaluation models were considered for comparison to the intuitive models developed by the Corporation. The two theoretical models that were chosen for this study were the Provus (1973) Discrepancy Evaluation Model and the Stufflebeam et al. (1973) C.I.P.P. Evaluation Model.

Discrepancy Evaluation Model

The Provus (1973) Discrepancy Evaluation Model was chosen specifically because it addresses evaluation as a "continuous information management process designed to serve as 'the watchdog of program management' and the handmaiden of administration in the management of program development through sound decision making" (Provus, 1973, p. 186). The Provus Discrepancy Evaluation Model suggests that evaluation is a process which should pass through five developmental stages: Definition, Installation, Process (interim - products), Product and Cost-benefit analysis (optional).

The Provus Discrepancy Evaluation Model is concerned with clearly defining objectives and then developing a plan to achieve those objectives. The Discrepancy Evaluation Model, simply put, examines the process of objective attainment and measures the gaps between what the intended outcomes and the actual outcomes are in order that changes may be made in the program to narrow the gap. This evaluation model is concerned with inputs, outputs, allocated resources and objectives. The emphasis is upon developing an operational plan which relies upon clearly defined specific objectives. For the preceding reasons, the Provus Discrepancy Evaluation Model was chosen as one of the two educational evaluation models to apply in a corporate setting.

C.I.P.P. Evaluation Model

The second model chosen was the Stufflebeam et al. (1973) C.I.P.P. Evaluation Model. This model is concerned with four developmental stages: Context, Input, Process and Product (hence the acronym C.I.P.P.).

The focus of the Stufflebeam et al. (1973) C.I.P.P. Evaluation Model, unlike Provus's objective oriented model, is towards making the evaluation model a management-like tool with the aim of assisting with the decision making process. Stufflebeam et al. (1973, p. 129) describe this model as "the process of delineating, obtaining, and

providing useful information for judging decision alternatives".

The literature on evaluation did not reveal studies which drew parallels between Provus's and Stufflebeam et al.'s theoretical models and business management concepts. This study has chosen the Provus and Stufflebeam et al. evaluation models to use in a comparison between theoretical evaluation models and intuitive evaluation models which are corporate based and are both field developed and tested. This comparison will address this study's question with respect to identifying some of the gaps between theory and practical application of educational evaluation models in a corporate setting in order to learn from these similarities and differences.

Origins of the Corporate Evaluation Committees

The evaluation committees within the corporation were struck for the purpose of evaluating the corporation's Claims Training Department and the corporation's litigation budgeting process. This study will first examine the origins of the Claims Training Evaluation Task Force Committee and then the Litigation Management Evaluation Task Force Committee. The work of these committees will be used as a vehicle to test the effectiveness of the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model in a corporate setting.

Claims Training Evaluation Task Force Committee

Initial discussions with the Claims Training Evaluation Task Force Committee chairman revealed that the corporation's Training Department had been receiving criticism about being out of touch with the needs in the field, that the method of centralized training was financially inappropriate given the regional nature of the corporation and that the training process itself was out of date given changes in adult education methodology and multimedia technology.

My field notes of May 5, 1994 captured the comments of a senior manager who was the corporation's Deputy Chief Underwriter. These comments reflect that there was concern with respect to core skill training and the use of technology within the corporation's education programme. A sampling of these comments is listed below:

- ... that claims training as it existed is out of touch with field
- the process of insurance claims has not changed since the 1930's and 40's, i.e., the way we do business has not changed
- technology (as) the kernel of insurance, has not changed
- emphasize that this is the message of the Chief Underwriter and the former Chief Underwriter.

Operating from this need to respond to the perceived concerns by staff and management, an evaluation committee was struck, under the sponsorship of a senior manager, to determine the future of the Claims Training Department.

Litigation Management Evaluation Task Force Committee

The Litigation Management Evaluation Task Force Committee was struck for the purpose of evaluating the corporation's litigation budgeting process. The budgeting process which was in place at the time of this study was established in order to control the expense of litigation by projecting the future costs of defending legal actions. Senior management, supported by data collected by the corporation's regional managers, felt that by accurately forecasting and assessing the cost of defending legal actions, these actions could be settled earlier for less expense by incorporating projected defence costs into the risk analysis process. Also, by incorporating a budgeting process for litigation, it was felt that defence counsel would be more likely to hold themselves accountable for providing an expected level of service rather than increasing corporate expenses through unnecessary delays and litigation inefficiencies.

Pre-evaluations Considerations

Senior managers in the corporation in each circumstance had to decide whether the evaluations would be conducted

either by an internal evaluation committee or by external consultants, or if it would be of value to conduct both an internal and external evaluation in order to have a complete, well-rounded evaluation. Senior management would have to take into consideration the advantages and disadvantages with using either approach or a combination of approaches. Some of these advantages and disadvantages include levels of bias, expertise (or lack of expertise) in the subject area being evaluated, time frames for completion of the evaluation and the overall cost of each of the various options available.

Internal Evaluation Advantages

With an internal evaluation, evaluators having an appreciation for the subject being evaluated could be seen as an advantage. Having a close familiarity with the subject matter should allow for greater insight. While there are elements of bias, when conducting internal evaluations there are advantages to be had from the interaction between evaluators who share a common interest.

As stated by Garaway (1995, p.87):

In participatory evaluation, the pool of participants tends to be a smaller group. While there is still pluralism, with the concomitant problem of conflicting interests, there is the added aspect of interaction which somewhat spreads the role of adjudication among the participants.

From my experience as a manager, utilizing internal staff for conducting evaluations is cost effective. Only in special circumstances would the purpose of an evaluation outweigh the expense of using external consultants. In other words, the situation would need to be of great importance for the cost of a dual evaluation to be warranted. Internal evaluations are cost-effective in that staff employees can be seconded to conduct evaluations and the cost to the corporation is limited to the loss of those employees' productivity. A wide range of points of view may be brought into such an evaluation by drawing upon the diverse nature of the corporation which means that the evaluators could be drawn from the large pool of employees available from any area of the corporation.

Again, as indicated by Garaway(1995, p. 98):

Participatory evaluation creates a learning environment in which evaluation findings are processed and accumulated by end-users in the very process of their being gathered.

In other words, an internal evaluation has the capacity to effect ongoing change as the evaluation unfolds which may be of considerable advantage to an organization.

Internal Evaluation Disadvantages

There are several disadvantages to utilizing an internal evaluation process. Stanley and Hopkins (1972, p. 4) stated:

The extent to which a measurement or evaluation is subjective is the degree to which personal bias and prejudice can influence scores. It is desirable to increase the objectivity of testing, interviewing, rating, and similar enterprises that often are pursued quite subjectively.

As the Barnes Craig & Associates report (1993, p. 6) stated:

Talleygrand is credited with saying, 'War is too serious a matter to be entrusted to the military.' So too, litigation has become far too serious a matter to be left in the hands of Lawyers.

This study considered the role of internal evaluators who were drawn from stakeholder groups which may have had vested interests in the outcome of the evaluation. For political reasons internal evaluations should be used carefully as personal and departmental agendas may influence the outcome of the evaluation. Even with the best of intentions and a lack of political influence, internal evaluators may simply be too close to the subject being evaluated and therefore blind to significant evaluation criteria. They may also lack evaluation expertise.

External Evaluation Advantages

There are advantages to using an external evaluation process. One advantage is that specific specialties of external evaluators could be selected in order to utilize their expertise in a particular area when such expertise is not available internally.

Outlined below are the services offered in the Barnes Craig & Associates report (1993, p. 1).

Licensed adjusting staff is drawn from both the legal community and insurance industry. This enables B.C. & A. to economically deliver [sic] high quality, innovative solutions by focusing the appropriate expertise on every project and case.

External evaluations, in contrast to internal evaluations, will generally provide an unbiased, uninfluenced perspective to the evaluation. In my experience, when an evaluation is being conducted in an area where the issues are volatile and emotional, an external evaluation process is able to provide an air of fairness and deflect hostilities from the parties who have commissioned the evaluation. In this sense, the evaluator partially assumes the role of mediator in the evaluation process. As indicated by Garaway (1995, p. 87):

In evaluation carried out by the participants, the external evaluator goes beyond being the primary investigator and participant observer to becoming a facilitator. As facilitator, his aim is to help transform a fairly natural process (evaluation being something we all undertake, personally, continually) into a broadly utilizable process.

External Evaluation Disadvantages

There are several disadvantages to an external evaluation. External evaluations have a tendency to be expensive and the results can sometimes be challenged on the basis of the external evaluators' apparently not being

familiar with the particular nuances of the object being evaluated. While the external evaluators hold no allegiance, their non-partisan role may be criticised on the basis that a lack of intimate knowledge of the evaluation object and its impact on the rest of the corporation may result in a misinterpretation of the data. This deficiency is identified in the Claims Training Project Proposal (April 6, 1994, p. 3) where the project team members are described as having "first hand experience of the information being generated and able to ask questions as appropriate".

A Combination Approach

A third option is to conduct both an internal and external evaluation to offset the deficiencies of both evaluation methods. This option is not usually viable due to the expense involved. Discussions with the Claims Training Evaluation Task Force Committee chairman resulted in the following summary from my notes on the strengths and weaknesses of a dual evaluation. These notes were not shown to the chairman for approval or comment.

In order to warrant the conducting of a dual evaluation, the issues must be serious and the potential results must have a wide-ranging impact on the Corporation. A dual evaluation will be both costly in terms of productivity and financing, but the results are usually very solid, with little opening for criticism if they are conducted in a competent fashion. The usual practice would be to conduct an internal evaluation initially and follow that by an external evaluation to confirm the results of the internal evaluation. The external evaluators can

draw upon the experience and findings of the internal evaluation once they have conducted their external evaluation in order to determine variances and explore the reasons for the variances. The dual evaluation also has the effect of acting as a buffer for management in that it will deflect criticism. (April 8, 1994)

Significance of the Study

This study examines two theoretical educational evaluation models, the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model, to determine what benefits, if any, could be had if they were applied in a corporate setting. By studying these models in a corporate setting weaknesses and strengths of the models will be highlighted. By studying these qualities, educational evaluators will be in a better position to improve upon the design of evaluation models. The evaluators in the corporate setting will be able to weigh the advantages of using a formal evaluation process to enhance corporate evaluations.

Significance to Education. Education would benefit from the knowledge gained of the practical needs and concerns of business. As well, a furthering of the understanding of theoretical models by an examination of any gaps that exist between the intuitive and theoretical models will be instructive. By studying these gaps educational evaluators will be in a position to enhance the practical application potential of evaluation models by avoiding

pitfalls and designing evaluation models which add value for the users of the models.

Significance to Business. The business world would benefit from the application of theoretical models, which may be tailored to specific needs, without having to "re-invent the wheel". As Robbins and Stuart-Kotze (1986, p. 24), state in their management theory text:

because theory and practice are often divergent, management is both a normative (what should be done) and descriptive (what is actually done) process. In other words, the practicality of the day-to-day job often takes priority over theoretical "how to's".

However, by taking theoretical evaluation models and learning about the similarities and gaps between what is expected to occur and what is actually occurring, relevant "pieces" may be sifted from the process in order to enhance the success potential of the evaluations that would otherwise be solely intuitive.

Two questions for consideration

1. Is there a need for this study?

In business as in education there is an ongoing need to improve efficiency and to forecast results accurately. As stated by Robbins and Stuart-Kotze, (1986, p. 179):

Organizations whose management can develop accurate forecasts of external and internal factors have a distinct advantage over their less successful competitors. If the variation management is attempting to predict follows some established pattern or relationship, there are forecasting techniques that can be valuable.

By looking at the variances between intuitive evaluation models and theoretical educational evaluation models it is possible for those differences and similarities to be explored and new lines of thought developed for educational evaluation purposes. The corporate based evaluation practices can be expected to show weaknesses in the educational evaluation models by virtue of the fact that gaps become apparent when applied to different populations and settings; in addition the education models may offer a perspective that has not been apparent to business.

2. How is the study delimited?

This study examines only two theoretical educational evaluation models and two intuitive corporate based evaluations.

This study has not explored the possible effects of an evaluation team upon the application of evaluation models, nor has it considered the possible influence of committee composition upon results.

While there is a strong business component to this study, there will be no discussion of business processes. It is not the intent of this study to develop strategies for business management.

Limitations of This Study

Two intuitive models were developed within a large

corporation and there could be variances among large corporations or between large corporations and smaller firms. Given different types of business, different types of evaluation processes may be required.

The participant observer role may create an element of bias in this study. If this study were to be replicated using an impartial observer, different observations may be made.

At times, my dual role as participant and observer was in conflict. As a participant in both committees, I had certain responsibilities which could have influenced my objectivity as an observer. In the Claims Training Evaluation Task Force Committee, I co-authored the final report and I acted as chairman of the presentation committee. As chairman, I led the presentation committee in delivering the results to the Manager of Human Resources Development. In this role I was expected to defend the work of the Claims Training Evaluation Task Force Committee and in particular the methodology, results and conclusions.

In the Litigation Management Evaluation Task Force Committee I authored the report and was a member of the presentation committee. In both cases, I had a personal and professional interest in the success of the committees' results. However, this conflict did not arise in either committee until after the data had been collected and

analyzed by the full membership of each committee. During the course of each evaluation, I was not aware of the role that I would take in authoring reports or presenting findings. The final reports of each committee were submitted for review by the committee members prior to being made public.

The rank and position of the Claims Training Evaluation Task Force Committee and Litigation Management Evaluation Task Force Committee members may have played a role in the functioning of each committee. Given different committee members with different status within the corporation, the results may have been different.

The time frames allowed for each committee may have influenced the approach taken by each committee. For example, the Claims Training Evaluation Task Force Committee was given approximately three months within which to conduct its evaluation of the Claims Training Department and present the results to senior management. The Litigation Management Evaluation Task Force Committee, on the other hand, was under no such time restriction. In fact it had no time restriction at all. Varying the time frames by either extending or restricting the time periods allowed for evaluation may have produced different results.

The method used for data collection was that of participant observer. The participant observer role may

have influenced the workings of both the Claims Training Evaluation Task Force Committee and the Litigation Management Evaluation Task Force Committee. Also, as participant observer, it was possible to move too easily from my role as participant observer into my role as an active participant on the committees.

There may have been an impact on the functioning of both committees by external personal and political influences. The impact of external influences will be discussed in Chapter Five. This factor is significant when examining evaluation programmes in corporate settings and is not taken into account in either the Provus Discrepancy Evaluation Model or the Stufflebeam et al. C.I.P.P. Evaluation Model. The potential is significant for corporate based evaluations to be affected by personal and corporate agendas and other external factors which may or may not be apparent to evaluation committees. This aspect underscores the whole of the evaluation process. It is a factor which potentially could put the members of an evaluation committee at risk and, as will be discussed, can apparently shape the focus of the evaluation process.

Summary

This study will use two theoretical educational evaluation models, the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model. These

models are applied retrospectively in a corporate setting in order to reveal strengths and weaknesses of the models. These strengths and weaknesses will be considered from the perspective of what can be added to the general knowledge of educational evaluation. In addition, these studies will be used to provide insight into enhancing evaluations conducted in a corporate setting.

CHAPTER TWO

LITERATURE ON EDUCATIONAL EVALUATION

Chapter Two will provide a brief overview of the history of Education Evaluation with a particular emphasis upon the Education Evaluation Models of the 1960's and 1970's. Models from this time period were chosen as it was the beginning of a new era in educational evaluation. This brief history is presented as a high level overview and is intended to provide general background information. The two models used in this study, the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model, will be discussed briefly in order to highlight their strengths and weaknesses. Other evaluation models will also be briefly outlined.

History of Evaluation in Education

Evaluation, as a component of the education process, was not prevalent until the latter half of the nineteenth century. According to Worthen and Sanders (1987, p. 12), evaluation in American education prior to the mid nineteenth century was essentially non-existent. The direction taken by education was influenced by religious or political policies as opposed to competency, need, etc. As stated in Worthen and Sanders (1987, p. 12):

Prior to 1837, political and religious beliefs dictated most education choices. Communities were

happy to attract and hold teachers, regardless of their competence, and if a teacher did prove incompetent in those days, formal evaluation was relatively pointless anyway - the school just closed for lack of students.

The foundation of data collection for educational purposes has been attributed to Henry Barnard, Horace Mann and William Torrey Harris. They worked for the state education departments of Massachusetts and Connecticut, as well as the United States Education Bureau, where they developed data collection processes for amassing information which could be used to make decisions regarding educational policy.

Worthen and Sanders (1987, p. 12) indicate that during the period 1838 to 1850, "Horace Mann submitted 12 annual reports to the Board of Education of the Commonwealth of Massachusetts". These reports contained concerns with respect to a number of areas, ranging from outside supervision to the selection or construction of curriculum materials. No mention was made as to how the data were collected, according to Worthen and Sanders.

During the latter portion of the nineteenth century, Joseph Rice conducted an assessment of large school systems across the United States. This assessment was conducted with the aim of establishing whether or not his theory of the inefficient use of school time was correct. As Worthen and Sanders (1987, p. 13) state:

He used these data to support his proposals for restructuring spelling instruction. His tests of arithmetic, on the other hand, revealed large differences among schools; consequently Rice proposed the setting up of standardized examinations.

Edward Lee Thorndike, at the beginning of the Twentieth Century, proposed that measuring developmental change was an important facet of the education process. Measurement took the form of testing, which became the accepted means of evaluating schools.

Guba and Lincoln (1989, p 24) make the following comments on the use of testing:

The utility of tests for school purposes was well recognized by leadership personnel. The National Education Association appointed a committee in 1904 to study the use of tests in classifying children and determining their progress; the association appointed three additional committees by 1911. In 1912, the first school district Bureau of Research was established in New York City.

Worthen and Sanders (1987, p. 13) state:

The testing movement was in full swing by 1918, with individual and group tests being developed for use in many educational and psychological decisions.

Though the early school system surveys had relied mainly on criterion-referenced tests to gather group information in school subject areas, the 1920's saw the emergence of norm-referenced tests developed for use in measuring individual performance levels.

In the forty-five year period between 1920 and 1965, greater emphasis and a focus upon the development of testing

processes occurred. As Worthen and Sanders (1987, p. 14) state:

The development of standardized achievement tests for use in large-scale testing programs was a natural outgrowth of this trend (State-wide testing). During this period, "evaluation" was most often used to mean the assigning of grades or summarizing of student performance on tests.

Ralph Tyler proposed, in 1932, an objectives-based approach to educational evaluation and developed measurement tools to support his concept. Wolf (1987, p. 3) observes, "In the 1930's, largely as a result of work by Ralph Tyler, evaluation was being formally conceptualized and a fledgling technology was developed." The objective-based approach involved defining objectives and examining outcomes to determine whether or not objective achievement was successful.

According to Worthen and Sanders (1987), the 1940's and 1950's were essentially a period where previous educational evaluation concepts were applied. Objectives for education were debated, agreed upon and implemented.

Throughout the mid 1950's and 1960's educational evaluation, which was based upon Tyler's approach, was further developed. Bloom et al. published in 1956, Taxonomy of Educational Objectives: Handbook I: Cognitive Domain. "Bloom's Taxonomy", is referenced by Worthen and Sanders (1987, p. 16) and focussed upon defining:

in explicit detail a hierarchy of thinking skills applicable to various content areas. This document continues to be a standard tool both in testing and in curriculum development, design and evaluation.

A major development in the United States in 1964 was the enactment of the Civil Rights Act which was followed by the passage of the Elementary and Secondary Education Act (ESEA) in 1965.

According to House, (1995, p. 15):

Prior to 1965, evaluation was a minor activity, a sideline engaged in by academics as extra consulting work. Then came the Great Society Legislation in the United States. With the passage of the Elementary and Secondary Education Act in 1965, everything changed. Senator Robert Kennedy insisted that an evaluation amendment be attached to the education bill, and evaluation became a federal mandate that spread to other social programs.

The effect of these acts was an outpouring of money into education to support education programmes for disadvantaged youths and educational research. As a control over this expenditure of funds, evaluation programmes became mandatory in order to hold educators and researchers accountable for how the funds were spent. As Worthen and Sanders (1987, p. 17) state:

Translated into operational terms, this meant that thousands of educators were for the first time required to spend their time evaluating their own efforts. Project evaluations mandated by state and federal governments have since become standard practice, with evaluation emerging as a political tool to control the expenditure of public funds.

The late 1960's was a transition period where educators had to develop extensive evaluation programmes in order to comply with the federal mandates of 1964 and 1965. Wolf (1987, p. 4) describes this period as one where:

The political popularity of the evaluation requirement quickly spread to other social legislation so, by the end of the 1960's, it was commonplace to require systematic annual evaluations of social programs.

Worthen and Sanders (1987, p. 17) identify a poignant quote of Guba (1967, p. 312) that comments upon the effectiveness of the evaluations that arose from the passing of this legislation.

None of these product evaluations will give the Federal Government the data it needs to review the general Title III program and to decide how the program might be reshaped to be more effective. (Guba, 1967, p. 312)

The above quote identified a grave concern that the then current evaluation programmes were inadequate to evaluate education programmes and provide meaningful data with which to satisfy the federal mandate. Operationally, there were no adequate guidelines for evaluators, which led to evaluations being created and conducted by inexperienced people. Wolf (1987, p. 4) further comments that:

The period from 1965 to the early 1970's was one of considerable confusion. A great deal of activity occurred under the heading of evaluation. Much of the work was highly questionable.

From 1967 to 1973 new strategies for evaluation were formulated, and several evaluation models were developed. Greater emphasis was placed upon evaluation, and evaluation as a field of study was, by necessity, created. With various evaluation models surfacing, debate arose as to the appropriateness and general application of these models. In 1967 the United States government created the Centre for the Study of Evaluation. This was followed, in 1972, by the creation of the National Institute of Education (NIE).

Worthen and Sanders (1987, p. 19) state that the:

NIE focused one of its research programs on evaluation in education, supporting field research that added to our knowledge of evaluation methodology, and also funded research to adapt methods and techniques from other disciplines for use in educational evaluation.

In the years following the creation of NIE, there has been a considerable amount of research and development of educational evaluation. Educational evaluation has become a field of study in its own right, and it continues to explore new methodologies and applications. As Worthen and Sanders (1987, p. 20) observe, educational evaluation:

must continue to grow and adapt to changing conditions and demands.

The resulting decrease in demand for evaluation of large scale, federally supported educational programs has led some commentators to make gloomy predictions about the future of evaluation in education and related areas.

The shift from large scale evaluations to evaluations at the local level is one method of adapting educational evaluation to meet the needs of local governments and school boards.

Wolf (1987, p. 5) summarizes the then current state of educational evaluation in contrast to the late 1960's to mid 1980's below.

Evaluation is now seen to be a more open and ongoing process intended to yield information that will lead to the improvement of educational programs. This latter view is in marked contrast to the view of 10-20 years ago when it was felt that the results of evaluation studies would be the most important determinant of continued support.

In my opinion, educational evaluation will continue to be an integral part of the "business of education" as long as the demands for responsible education and judicious use of public funds remain political and public concerns.

Education Evaluation Models Used in This Study

As stated in Chapter One, the two evaluation models chosen for this study are the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model. Each model has advantages as well as limitations.

Provus Discrepancy Evaluation Model

Advantages. The Provus Discrepancy Evaluation Model has the advantage of using a straightforward concept which is basic to evaluation. That is, it clearly identifies what

it is that will be evaluated by concentrating upon defining objectives. The balance of the evaluation is then focussed upon a comparison of actual outcomes with stated objectives. This is a very clean concept that is easy to follow and should produce definitive results. This overview oversimplifies the actual model, but does capture the essence of the model's major advantages.

Limitations. The Provus Discrepancy Evaluation Model is limited in that it has been criticized for being too narrow in focus. Worthen and Sanders, (1987, p. 73), indicate that evaluation models such as the Provus Discrepancy Evaluation Model with an objectives oriented approach have their critics. They summarize these concerns in nine points as listed below.

The objectives-oriented evaluation approach:

- 1) lacks a real evaluative component (facilitating measurement and assessment of objectives rather than resulting in explicit judgements of merit or worth).
- 2) lacks standards to judge the importance of observed discrepancies between objectives and performance levels.
- 3) Neglects the value of the objectives themselves.
- 4) Ignores important alternatives that should be considered in planning an educational program.
- 5) Neglects transactions that occur within the program or activity being evaluated.
- 6) Neglects the context in which the evaluation takes place.

- 7) Ignores important outcomes other than those covered by the objectives (the unintended outcomes of the activity).
- 8) Omits evidence of programme value not reflected in its own objectives.
- 9) Promotes a linear, inflexible approach to evaluation.

Given Worthen and Sanders' (1987) comments, it is apparent that this type of evaluation model has severe limitations which could affect the value of the results obtained. In my study, the Provus Discrepancy Evaluation Model is considered from the perspective of its application in a business setting. What may be seen as limitations for use in education may have practical advantages when used in an environment where the focus is generally upon objective attainment.

Stufflebeam et al. C.I.P.P. Evaluation Model

Advantages. The Stufflebeam et al. (1972) C.I.P.P. Evaluation Model is a management-oriented evaluation approach. Worthen and Sanders (1987) indicate that this approach is designed to serve the needs of management in the decision making process. The focus of this approach is such that evaluation can start at the outset of a programme and provide ongoing information which will aid in the development of the programme. According to Worthen and Sanders (1987), this approach takes advantage of

opportunities as they arise and allows management to make informed decisions at the time those decisions need to be made.

A particular advantage of using the Stufflebeam et al. C.I.P.P. Evaluation Model is that it places the objectives of an evaluation in context. This allows for a more complex evaluation to be conducted. That is, relevant data may be collected to support questions which have a greater complexity.

As Worthen and Sanders (1987, p. 84) state:

The C.I.P.P. Model, in particular, is a simple heuristic tool that helps the evaluator generate potentially important questions to be addressed in an evaluation.

The management-oriented approach to evaluation supports evaluation of every component of an educational program as it operates, grows, or changes. It stresses the timely use of feedback by decision-makers so that education is not left to flounder or proceed unaffected by updated knowledge about needs, resources, new developments in education, the realities of day to day operations, or the consequences of providing education in any given way.

An example of this process occurred in the Litigation Management Evaluation Task Force Committee proceedings. The committee ascertained that there was no compliance with Bulletin CDB894 and then used that information to refocus its mandate. By refocussing its mandate, the Litigation Management Evaluation Task Force assumed ownership of the development of a new litigation management strategy.

Limitations. This type of management-oriented evaluation model is limited in that it serves the needs of the decision makers and may restrict or impede the evaluator's exploration of other issues that arise through the course of the evaluation. While some of these potential issues may be important, they will be overlooked in favour of complying with the objectives and directions of the decision makers.

This leads to a second weakness in that this type of evaluation may be subject to political or personal agendas which could shape the outcome of an evaluation. There is concern in this study that this variable is active due to the presence of political and personal influences within and surrounding both the Claims Training Evaluation Task Force Committee and Litigation Management Evaluation Task Force Committee. This variable was not specifically tested and would therefore act as a limitation on the generalizability of this study. This issue will be dealt with further in Chapter Three.

Another limitation is the cost factor related to conducting an evaluation of this type in its entirety. As Worthen and Sanders (1987, p. 85) state, "if followed in its entirety, the management-oriented approach can result in costly and complex evaluations".

Other Education Evaluation Models Considered

The following educational evaluation models were considered for use in this study.

1. Tylerian Evaluation Approach
2. Hammond's Evaluation Approach
3. U.C.L.A. Evaluation Model
4. Stake's Countenance Model

Each model will be briefly discussed in order to provide a background of alternative evaluation models.

Tylerian Evaluation Approach. Ralph Tyler, during the late 1930's and early 1940's, developed an educational evaluation model. This education evaluation model was intended to look at programme objectives and determine whether or not they had been attained.

This model is very linear, which makes it simple to understand and apply. It consists of seven basic steps. Worthen and Sanders (1987, p. 63) identify these seven steps below:

1. Establish broad goals or objectives.
2. Classify the goals or objectives.
3. Define objectives in behavioural terms.
4. Find situations in which achievement of objectives can be shown.
5. Develop or select measurement techniques.
6. Collect performance data.

7. Compare performance data with behaviourally stated objectives.

In comparing objectives with performance outcomes, it is possible to determine the success or failure of the programme. Early identification of variances between objectives and performance outcomes allows for ongoing modifications to be made in order to enhance the potential for programme success.

Hammond's Evaluation Approach. R. Hammond (1973)

developed an evaluation approach which followed the Tylerian approach very closely with one exception/innovation.

Hammond's evaluation model consists of six steps. Worthen and Sanders (1987, p. 68) identify these steps as listed below.

1. Defining the program.
2. Defining the descriptive variables (using his cube).
3. Stating objectives.
4. Assessing performance.
5. Analyzing results.
6. Comparing results with objectives.

The major difference between Hammond's Evaluation Model and the Tylerian Approach is that Hammond has added a third dimension to the evaluation approach. Hammond's cube (as

noted in step 2 above) is a tool which may be used by the evaluator to generate a number of questions that may be explored in the evaluation. In essence, it consists of three dimensions (instruction, institution and behavioural objectives) which are broken down into smaller divisions. Where these smaller divisions intersect within the three dimensional cube relational questions are generated or suggested. If all factors apply to a given evaluation, a maximum of ninety cells are available to generate questions about relationships. The number of factors available are reduced in each dimension in accordance with their applicability to the evaluation undertaken. Worthen and Sanders (1987, p. 68) state that the value of Hammond's evaluation approach is as "a valuable heuristic tool the evaluator can use in analyzing the successes and failures of an educational activity in achieving its objectives".

U.C.L.A. Evaluation Model. M.C. Alkin (1969)

developed an evaluation model that closely parallels the Stufflebeam et al. (1973) C.I.P.P. Evaluation Model. Worthen and Sanders (1987, p. 81) make the following observation.

Both the C.I.P.P. and U.C.L.A. frameworks for evaluation appear to be sequential, but the developers have stressed that such is not the case. For example, the evaluator would not have to complete an input evaluation or a systems assessment in order to undertake one of the other types of evaluation listed in the framework.

While the application of the U.C.L.A. Evaluation Model may not necessarily be linear, the actual structure is essentially the same as the Stufflebeam et al. C.I.P.P. Evaluation Model. Table 1 presents a comparison of the stages of both the U.C.L.A. and Stufflebeam et al. C.I.P.P. Evaluation Models.

Table 1

U.C.L.A. and C.I.P.P. Evaluation Models - Comparative Processes

<u>U.C.L.A.</u>		<u>C.I.P.P.</u>	
Stage	Process	Stage	Process
1	Systems assessment	1	Context evaluation
2	Program planning	2	Input evaluation
3	Program implementation		
4	Program improvement	3	Process evaluation
5	Program certification	4	Product evaluation

The basic difference between the models is that the U.C.L.A. Evaluation Model does not require one stage to be completed before passing on to the next. It allows the evaluator to "cycle" from one stage to another, continually revising and reviewing as the need arises.

Stake's Countenance Evaluation Model. R.E. Stake (1967) developed his concept that evaluations should consist

of two basic components, "Description" and "Judgement".

These two components were proposed as the two countenances of evaluation. The purpose of this model was to provide the evaluator with a tool for organizing and analyzing data.

Worthen and Sanders (1987, p. 130) describe the workings of the Stake's Countenance Model below.

The evaluator would analyze information in the description matrix by looking at the congruence between intents and observations, and by looking at dependencies (contingencies) of outcomes on transactions and antecedents, and of transactions on antecedents. Judgements would be made by applying standards to the descriptive data.

Stake's Countenance Model provides the evaluator with a conceptual framework as opposed to an evaluation formula.

Rationale for Selecting the Provus and Stufflebeam et al.

Evaluation Models

The Provus and Stufflebeam et al. evaluation models were chosen for this study for the following reason. I wanted to use evaluation models that were not complex in structure in order to keep the study simple. The more dimensions that are added to this study the more complex it could become and this might cloud some basic issues. The Tyler model was too linear and the balance of the models considered added too many dimensions for this study to consider. The Provus and Stufflebeam et al. evaluation models seemed to fit in well with the corporate environment

and both of these models seemed to be capable of providing information that could be of value in a corporate setting.

Education Evaluation Research Sources

Several sources were consulted to determine if similar studies to mine exist and to identify critiques of the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model.

Aside from direct access to the textual material, as listed in the bibliography, an internet search was conducted. This search used Yahoo as the search engine and was conducted through the U.B.C. Library Web Page.

The search parameters were restricted to the following key words and phrases.

1. Provus
2. Stufflebeam
3. Discrepancy Evaluation
4. C.I.P.P.
5. Evaluation
6. Education
7. Education Evaluation.

In total, one hundred and twenty-four records were searched and hits occurred in one hundred and thirteen of these records. On examination of the one hundred and thirteen hits, there were no instances of studies which were of a similar nature to my thesis. There were six hits under

the Provus search and three hits under Stufflebeam. None of these nine hits appeared from their synopses to critique the models or their applications, but were instances where the models were referenced in relation to other studies.

Only one hit was achieved under Discrepancy Evaluation. This hit was Discrepancy Evaluation For Educational Program Improvement and Assessment, (1928) [sic] by Malcolm Provus.

Results of Documentation Search

While the search I conducted was by no means exhaustive, it is reasonable to conclude that this study appears to be unique in its focus.

Summary

Chapter Two has presented a review of the history of educational evaluation and has discussed the models used in this study. Some potential alternative models were also discussed and potential sources of further information through an internet search were presented.

CHAPTER THREE

METHOD

Chapter Three will discuss the methodology used in this study. The methodology will be presented from two perspectives. The first perspective will examine the methodologies used by both the Claims Training Evaluation Task Force Committee and the Litigation Management Evaluation Task Force Committee. The second perspective will present the method of data collection that I used in order to allow for potential replication of this study. Chapter Three will also discuss the makeup of the committees involved in this study and my role as a participant observer. The methodologies used by both committees will be presented and related to the two educational models chosen for this study, the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model.

Task Force Members

Two evaluation task forces within a large corporation, the Claims Training and the Litigation Management Evaluation Task Force Committees, were examined. Eleven members (six managers and five field staff) and a secretary, who would act as recorder only, comprised the Claims Training Evaluation Task Force Committee. The nine member Litigation Management Evaluation Task Force Committee was comprised of

five claims managers, one claims examiner, three lawyers, and a clerk who would act as recorder only.

Role of the Participant Observer

In both committees I functioned as a participant observer and did not take an active role in the actual design of the evaluation, but participated in their activities. Both committees were advised in their initial meetings that I would be acting as a participant observer for the development of my thesis and that I would also be a contributing member of each committee.

As a participant in both committees I was expected to take an active role in the workings of the committees. In the Claims Training Evaluation Task Force Committee I co-authored the final report and in the Litigation Management Evaluation Task Force Committee I authored the final report. I acted as chairman of the presentation committee for the Claims Training Evaluation Task Force Committee. As chairman I led the presentation team in delivering the results to the Manager of Human Resources Development. In this role I was expected to defend the work of the Claims Training Evaluation Task Force Committee and in particular the methodology, results and conclusions. I was a member of the presentation team Litigation Management Evaluation Task Force Committee, but I did not assume a lead role.

The guidelines for the operation of each committee varied slightly, but it was understood by all committee members that title or rank within the corporation did not play a role in the committee. For example, at the initial meeting of the Litigation Management Evaluation Task Force Committee on October 18, 1993 the Chairman of the committee set out the "Rules of Conduct" as follows:

- 1) Rank has no privilege. We are all equal partners.
- 2) No idea is a bad idea.
- 3) All suggestions will be approached with an open mind. Only constructive criticism will be allowed.
- 4) We will remain neutral in the face of vested interests.
- 5) Creativity will be respected.
- 6) Decision making by consensus.
- 7) All are equally responsible for the success of the project.

Participants in the Claims Training Evaluation Task Force Committee Investigation

The Claims Training Evaluation Task Force Committee operated under the sponsorship of a senior manager who commissioned the project. The Task Force team members consisted of an office manager, a claims manager (the participant observer), and a regional manager as field

management representatives. Also representing management was a head office support manager, an organization development manager and a systems manager who would act as facilitator. The five staff members who made up the balance of the committee represented a cross section of field staff members who were directly impacted by Claims Training. A secretary was also involved as record keeper.

The focus groups and individuals interviewed were selected by the Claims Training Evaluation Task Force Committee to ensure cross divisional, regional and geographic representation. They were chosen to represent each of the interested groups who were affected by the training which was conducted by the Claims Training Department.

Criteria For Focus Group Selection

The Claims Training Evaluation Task Force Committee decided that there would be one focus group chosen as a representative of each work group affected by the Claims Training Department. Each focus group consisted of 15 members and each selected member was sent an E-mail which outlined, in general terms, what the purpose of the committee was, but did not provide specific questions or guidelines with respect to what would be asked in each focus group session. The focus groups were balanced for gender, regional representation and length of service.

Each focus group was asked the same questions for consistency between groups. The members of the focus groups were seated in a conference room and asked specifically prepared questions in a roundtable format. Their answers were captured by individual committee members in private notes and by the committee secretary. The secretary's notes were documented in MS Word, printed and projected on an overhead projector for the participants to view during each focus group session. Each focus group session was held to a two hour interview period and each group was interviewed only once.

Task Force Procedures

The project team chairman announced in the first meeting of the Claims Training Evaluation Task Force Committee on April 8, 1994 that the best method of obtaining an accurate assessment of the claims training situation was to review the historical documentation which existed regarding previous assessments of the Claims Training Department and the recommendations that were made. No other options were discussed and the Claims Training Evaluation Task Force Committee Chairman set the direction for the committee. Following the historical review, the project team received a series of presentations by managers who either had a vested interest in the future of the Claims Training Department, or who were affected directly by the

product of the Claims Training Department. In addition, the project team interviewed the recipients of training by the Claims Training Department by way of focus group sessions. The presentations and focus group sessions were conducted within the context of the three objectives prescribed in senior management's mandate for the Claims Training Evaluation Task Force Committee. The three objectives set out by senior management are listed below.

1. To determine the best organizational structure for providing the training.
2. To determine what claims training is needed.
3. To determine who could best deliver what claims training.

Methodology: Claims Training Evaluation Task Force
Committee - Historical Background

Six documents were reviewed by the Claims Training Evaluation Task Force Committee prior to the commencement of the presentations and interviews. These documents were selected by the chairman of the Task Force for review by its members in order to provide an historical perspective. The documents reviewed by this Task Force were:

1. Claims Manpower Planning: Training Sub-committee, January 1991.
2. Strategy for Claims Training Redevelopment: Gateway Systems Services, June 1991.
3. Claims Task Force on Training & Development: Report No. 1, August 1991.

4. Claims Task Force on Training & Development: Final Report, November 1991.
5. Joint Task Force Report on Claims Issues, February 1992.
6. Corporate Training Consortium Initiative, September 1993.

Methodology: Claims Training Evaluation Task Force

Committee - Current Situation

In total, the project team heard presentations from ten managers who were directly involved with claims training at a departmental level and interviewed approximately eighty staff and managers through the focus group process.

The ten managers who were involved at a departmental level were chosen for their specific direct involvement in Claims Training. Their positions placed them in a role of either being responsible for the delivery of effective training (not the actual delivery) or for the support of Claims Training through the provision of resources or planning strategies. For example, corporate financial results may be dependent upon the effectiveness and quality of the training received by staff.

The importance of effective training on claims severities (dollars spent per claim) was summed up by one senior manager. Given his position as a senior manager, his comment carried a great deal of weight with the committee

and could have influenced the committee's perspective. The senior manager, (May 5, 1994) stated:

Severity control (controlling the number of dollars spent per claim) important, but to dwell on it is wrong focus. If you do it properly, you will get severity under control. Lot of pressure on us. First three months of this year were a disaster. April was good.

From this perspective, it was not the training itself that was important but rather the outcome of training. This is typical of the perspectives of the ten managers involved in claims training at a departmental level. These managers were results oriented rather than process oriented. These managers were not focussed upon the "how's" and the "why's" of training, but rather on the results. Their concerns and interests come from a corporate or macro perspective, where their focus is upon results. They were not concerned with managing at a micro level which involves the actual workings of claims training. From their viewpoint, they were able to see the effect of claims training upon the corporation's financial picture and their concerns would have a different emphasis than the concerns of those people involved with the delivery of claims training.

In contrast to the presentations made by the ten managers who were involved in claims training at a departmental level, a sample cross section of the recipients of training by the Claims Training Department was

interviewed through the focus group process. As recipients of training it was expected that they would be more process oriented than results oriented as they were directly impacted by the delivery of training. The data obtained were reviewed upon collection after each focus group/interview session by the entire Claims Training Evaluation Task Force Committee in order to ensure that by examining the data for trends, the weaknesses or strengths of the current state of claims training could be immediately identified. Upon conclusion of the data collection process for each group, the data were analyzed and general issues and specific issues were identified by the Claims Training Evaluation Task Force Committee which led to a series of recommendations and supporting rationale being developed.

The data were collected between May 6 and May 20, 1994. This evaluation process was very focussed and the Claims Training Evaluation Task Force Committee was very cognizant of its time-limited mandate. The project team was required to make a presentation of its findings to senior management by June, 1994.

The presentations to the Claims Training Evaluation Task Force Committee by the ten managers who were directly involved with Claims Training at a departmental level were organized in such a fashion that the schedule could be completed within the fewest number of days possible. The

presentations were completed within two sessions. The process of capturing information involved personal note-taking, as well as note-taking by the secretary on a laptop computer. The notes from the laptop computer were printed and distributed to the Claims Training Evaluation Task Force Committee members. These notes were then compared with the personal notes made by the committee members at the time of the interviews. This was an involved process but the level of detail was considered important and necessary by the committee to ensure accuracy. The committee as a whole amended or corrected the notes on the laptop computer to create the "official" notes immediately upon conclusion of each presentation. It was felt that this method would allow the project team to collect accurate information by reviewing the presentation in as short a time-frame as possible after the presentation was made. This method of note-taking was used on a trial basis during the initial presentations and upon completion of the presentations received endorsement from the project team for use in the focus group presentations as well as all further meetings.

As a participant observer, I maintained personal notes and contributed those notes to the group session, which ultimately formed part of the formal project team notes. All members of the project team were aware that I was participating in this process as a participant observer, but

that had no apparent influence upon the functioning of the committee. I made this assumption based on the fact that except for the initial meeting, the subject of my role in the study was never discussed again by any of the committee members. The flow of each meeting did not appear to be hampered by my presence which led me to believe that the perception of my presence on the committee was focussed upon what I could contribute as opposed to what I was observing. My role as participant observer posed no threat to the members of the committee as the results of my study would have no direct influence upon the outcome of the evaluation nor, on a more personal level, the future of their careers. I attended all focus group sessions with the exception of two focus group sessions which were held concurrent with sessions that I was attending.

Litigation Management Evaluation Task Force Committee

The Litigation Management Evaluation Task Force Committee functioned in a different manner from that of the Claims Training Task Force. The Litigation Management Evaluation Task Force Committee was concerned with an evaluation of the corporation's handling of the litigation budgeting process as outlined in an internal directive, Bulletin CDB894. The project team for this task force was specifically chosen. The Litigation Management Evaluation Task Force Committee, in a similar fashion to the Claims

Training Evaluation Task Force Committee, operated under the sponsorship of a senior manager. Unlike the Claims Training Evaluation Task Force Committee, the Litigation Management Evaluation Task Force Committee was under no strict time constraint. The emphasis was to be on a quality evaluation with the outcome being practical solutions to the perceived problem. The Litigation Management Evaluation Task Force Committee operated under loose restrictions. The Committee began on October 18, 1993 and culminated in June 1996 with the rollout of a corporate litigation management strategy programme. This document had the "blessing" of the office of the Attorney General. The result was the development of new guidelines for the handling of procedures, relationships and the creation of a new reporting format. The original mandate to evaluate Bulletin CDB894 consequently resulted in its replacement altogether with an entirely new direction in corporate focus.

The chairman of the Litigation Management Evaluation Task Force Committee advised me in a private conversation near the beginning of the project that he was selected by senior management for his expertise in litigation claims handling. He also advised that the members of the Litigation Management Evaluation Task Force Committee were specifically selected by the Litigation Management Evaluation Task Force Committee chairman on the basis of

having a known expertise in the field of litigation claims handling. The project team was made up of one office manager as the chairman, a head office claims manager (specializing in large dollar, high profile claims), five claims managers (one of which was the participant observer), one claims examiner, three lawyers and a clerk. These members were selected on the basis of having been recognized as people with strong litigation backgrounds. Assisting the committee was a secretary and the Claims training representative would act as facilitator. (I would reprise my role as participant observer under the same conditions as in the Claims Training Evaluation Task Force Committee.) The Litigation Management Evaluation Task Force Committee by virtue of its composition was a very powerful committee. It combined expertise and experience in such a manner that the composition of the committee could have had an effect on the nature of the results and the development of the evaluation process. In other words, the positional power and specialized knowledge of the committee members may have been an influencing factor upon the evaluation. However, with respect to the original mandate of the committee, which was to determine whether or not there was compliance with Bulletin CDB894, the evaluation required a simple "yes or no" response from the subject group. The composition of the committee, in this instance, would not likely have affected

this outcome but may have played a role in the future workings of the committee.

The Litigation Management Committee first met on October 18, 1993 to discuss the scope of the project and develop a methodology which would be suitable for fulfilling the mandate. This resulted in a redefining of the mandate and clarified objectives. From the Litigation Management Evaluation Task Force Committee minutes of November 1, 1993, item number five states:

It was agreed that defence counsel must have a clear understanding of what the Corporation's expectations are for handling litigated files. Such expectations should also encompass the adjusters' duties. It was felt that the expectations could be presented in the form of a contract and will be addressed as a separate issue, but in conjunction with the reworking of CDB894.

The above quote was the first indication that the Litigation Management Evaluation Task Force Committee had changed its focus from simply determining whether or not there was compliance with Bulletin CDB894 to developing a new litigation management strategy.

Having redefined its purpose, the mandate and objectives, the Litigation Management Evaluation Task Force Committee's next task was to develop an evaluation methodology which would assist the committee in achieving its new mandate.

The Litigation Management Committee decided that the best method of obtaining an accurate assessment of the state of the current litigation management practices was to review Bulletin CDB894 and interview staff members and defence counsel. The committee also conducted a review of the literature of the reference material available on the issue of litigation management. Defence counsel and staff members were asked to respond to the following questions listed below.

1. Do you comply with Bulletin 894?
2. What about Bulletin 894 works or doesn't work?
3. Are the requirements of Bulletin 894 necessary?
If not, what are the alternatives?
4. What can be done to improve Bulletin 894?

Seven members of the Litigation Management Evaluation Task Force Committee were selected to interview seven senior lawyers who acted as defence counsel for the corporation. The seven senior lawyers were individually interviewed in one-on-one half hour sessions. Each interview was conducted by one committee member interviewing one lawyer. Each of the seven interviews was conducted by a different committee member. The lawyers were asked the four questions listed above. It was felt that by interviewing senior counsel, the committee would be able to take advantage of both their

experience and expertise with the litigation management process.

In total, forty-nine documents/reports were reviewed and the committee conducted individual interviews with seven defence counsel and group interviews with approximately fifty staff members. The data were obtained between October 18, 1993 and November 4, 1993 and discussed within two weeks of collection. The results of this data collection process were used immediately as part of the formative evaluation of the committee's mandate. In other words, the Litigation Management Evaluation Task Force Committee used the data upon collection to establish whether or not there was compliance with Bulletin CDB894. This allowed the committee to develop a "picture" of what the final results would look like and establish an understanding so that when the last of the data were collected, the conclusion was already evident. The committee was therefore able to broaden the scope of its investigation in order to make informed recommendations with regard to future litigation management procedures and defence counsel expectations.

During the course of the committee's investigation into litigation file management, the committee was asked to develop "canned defence plans". As a result, a sub-committee was struck and a defence plan for fibromyalgia (a non-specific soft tissue injury condition diagnosed by pain

response to 12 or more of 18 "trigger points". A confirmed diagnosis of fibromyalgia could command a large monetary award in court) was developed. In conjunction with the development of sub-committees, a head injury committee was also formed as part of the Litigation Management Evaluation Task Force Committee's mandate. The members of the Litigation Management Evaluation Task Force Committee, including me, were not directly involved with these sub-committees.

The result of the evaluation process was that the original mandate of the committee was refocussed into an examination of the entire litigation management process rather than the original narrow focussed examination of Bulletin CDB894.

The methodologies used in both task forces are representative of two evaluation processes, the review of the literature and the focus group evaluation. The review of the literature involved reading articles that were written by external sources which were available to the public. As well, the review involved an examination of internal reports which had limited circulation within the corporation. These reports were not generally accessible within the corporation, nor were they available to the public. The essence of the focus group evaluation process is to conduct personal interviews that, as stated by Worthen

and Sanders, (1987, p. 108), "allows clarification and probing". Interviewing is helpful in determining, by careful assessment, the values and beliefs of the people who are directly affected by the outcome of the evaluation. The review of the literature approach consists of an exhaustive search of all relevant documents pertaining to the specific subject area and then interviewing people who are directly affected by the process against the background of the historical documentation. The difference between the two approaches is that the former seeks to determine the current need whereas the latter seeks to place the current situation within an historical context.

Data Collection

The method used for data collection was that of participant observer. Acting in the role of participant observer had disadvantages and limitations. As Cousins (1996, p. 20) observed:

The most significant contribution of the present study is its evidence that, indeed, the level of researcher involvement in participatory evaluation does make a difference. Specifically, while the experience was generally positive for the school-based research committee members, the researcher's full partnership role may have led to the establishment of unrealistic expectations of the committee and its report.

Everyone involved in both the Claims Training Evaluation Task Force and Litigation Management Evaluation Task Force Committees was aware of my role as participant

observer and there was a short transition period which had to be undergone before the group was comfortable with that role. For example, the simple act of taking notes could make some members feel uneasy. The committee members could be apprehensive in not knowing whether the notes being taken were for the purposes of the committee or the study. This settling-in period was short-lived in this study as the committees became more involved and focussed on the evaluations themselves. As evidence for this, I observe from my field notes that there was no reference to my role as a participant observer after the first meeting of the committees.

Consequently, the participant observer aspect, in this study, apparently faded into the background very quickly in the minds of the other evaluators. As a participant observer I had to be aware that it was very easy to focus on the participation role in the evaluation rather than on the observation role, and the data could be skewed due to a lack of clear focus. It was a delicate balance. As Cousins(1996, p. 23) suggests:

If participatory evaluation is to become a viable approach to supporting organizational decision-making processes and enhancing organizational learning capacity, a much more realistic perception of what is entailed is needed.

Cousins(1996, p. 23) raises an interesting question with respect to the trade-off between quality and quick results

with respect to supporting participatory evaluation and the participant observer when he asks:

Will dissatisfaction with lame results from technically inferior "quick and dirty" studies adversely affect such attraction or enhance administrators' propensity to respect more technically sound and costly projects?

The simple answer could be that whether or not an organization chooses to use participatory evaluation and the participant observer will largely depend upon the organization making a business decision with respect to the needs of the organization at the time that the evaluation is commissioned.

In the Claims Training Evaluation Task Force Committee, I collected data by way of taking extensive personal notes and assembling copies of all documents (printed and electronic). The notes were intended to capture my observations and the main points being discussed in wording as close as possible to what was being used. The balance of the documents was collected with the intent of having a copy of everything generated by the Claims Training Evaluation Task Force Committee (including drafts) in order to create an accurate body of documentation.

In the Litigation Management Evaluation Task Force Committee the data collection process was different. It consisted mostly of a collection of documents with an

emphasis upon electronic e-mails, minutes and notes. Some personal notes were taken, but these were not extensive. This Litigation Management Evaluation Task Force Committee met irregularly, and much of the work of the committee was conducted by committee members in isolation. The meetings were used to update committee members on the status of various individual tasks.

Summary

A brief history of educational evaluation was discussed in order to provide the reader with an historical context within which to place the development of the educational evaluation models used in this study.

There are certain aspects of each task force's methodology which seem to parallel the theoretical models of Provus and Stufflebeam et al. very closely. The methodologies used by the two corporate task forces were outlined as well as the methodology used by the participant observer for data collection. Chapter Four will present the data collected in this study which will be used to explore the relationships between the theoretical evaluation models of Provus and Stufflebeam et al. and the two corporate based evaluations. The discussion and interpretation of the results will be presented in Chapter Five.

CHAPTER FOUR

RESULTS

The two corporate evaluations examined in the study, the Claims Training Evaluation Task Force Committee and the Litigation Management Evaluation Task Force Committee, were analyzed using two models designed for the evaluation of curriculum: the Provus (1973) Discrepancy Evaluation Model and the Stufflebeam et al. (1973) C.I.P.P. Evaluation Model. In addition, the records of the projects were explored from the point of view of the influence of external factors. From the perspective of the evaluation models themselves, comparisons were made to illustrate how the practical application of each model may affect the outcome of evaluations.

The two educational evaluation models were applied retrospectively to two evaluations that were conducted in a corporate setting in order to explore, hypothetically, how these evaluations may have benefited from the application of such models during the evaluation process. Neither project conducted its evaluations by utilizing a formal evaluation process. The effect of factors outside the evaluation models was also examined to determine possible influences upon both the application of the models and the

interpretation of the data collected. The data collected in this study are presented from both perspectives.

As will be discussed below, The Provus (1973) Discrepancy Evaluation Model was helpful for assessing approaches taken by two evaluation committees and the impact that good planning or the lack of it has on the quality and usefulness of evaluations. The Stufflebeam et al. (1973) C.I.P.P. Evaluation Model adds to Provus's focus by examining the context in which the evaluations take place. However, both models failed to account for the influence that external factors may have upon evaluations. As well as the effect of external factors, there is no indication by either Provus or Stufflebeam et al. that any consideration has been given to the possibility that the application of either model may not rigidly follow the evaluation guidelines exactly as set out due to external factors.

Evaluation Model Application Comparison

The Provus (1973) Discrepancy Evaluation Model consists of five stages.

Stage I - Programme Design. The first stage of the Provus Discrepancy Evaluation Model, defining the problem to be studied in terms of the programme design, was helpful in understanding the dynamics of the Litigation Management Evaluation Task Force Committee and the importance of clearly identifying programme objectives. The Provus

Discrepancy Evaluation Model was less helpful in the Claims Training Evaluation Task Force Committee, where there was a conspicuous lack of clearly defined objectives.

Nonetheless, The Provus Discrepancy Evaluation Model was instructive because of the contrast between the absence of clear objectives in the Claims Training Evaluation Task Force Committee with the theoretical need for clear objectives as outlined in the Provus Discrepancy Evaluation Model.

Stage I of the Provus (1973) Discrepancy Evaluation Model requires that the programme design be clearly laid out in terms of three dimensions: Input, Process and Output. This first stage, according to Provus, is crucial in creating a foundation for the evaluation in that it is necessary to identify what the intended objectives are, the plan for achieving those objectives and what success should look like if those objectives are achieved. Provus suggested that in Stage I the analysis of the problem is paramount in that without a properly defined problem there is an unclear reference upon which to base the remainder of the analysis.

For the Claims Training Evaluation Task Force Committee, the first stage of the Provus Discrepancy Evaluation Model suggested that the committee would find some difficulty completing its task effectively. This

prediction was made based on the discrepancy between the need for clearly defined objectives, as per the Provus Discrepancy Evaluation Model, and the lack of clearly defined objectives as was evidenced in the Claims Training Evaluation Task Force Committee.

Although a mandate to evaluate claims training was given to the Claims Training Evaluation Task Force Committee by senior management, this mandate did not provide specific objectives with respect to evaluating the claims training programme. The mandate was given to the Claims Training Evaluation Task Force Committee to provide general recommendations to the Claims Division and the Human Resources Division concerning the future of claims training. Rather than evaluate Claims Training within the scope of its performance, the recommendations were to focus on the clients' needs and specify the following:

1. How claims training needs can best be addressed.
2. What claims training is needed.
3. Who should deliver the claims training.

These objectives could be used to determine the marketability of the then-current programme but added little to establishing whether or not the Claims Training Programme had performed effectively. The actual intended performance outcomes of the Claims Training Programme were not stated.

The broad mandate was neither questioned nor further defined by the Claims Training Evaluation Task Force Committee. The Committee spent no time determining what the mandate specifically meant with respect to the performance of the Claims Training Department, but rather focussed on the product being delivered. In the first meeting, according to the Provus Discrepancy Evaluation Model, it would be expected that the problem would be identified, clarified, or at least discussed and a plan of action developed with respect to how the evaluation would be conducted. Instead, at the first meeting of the Claims Training Evaluation Task Force Committee, a great deal of time was spent discussing how changes would affect the members of the Task Force Committee who were directly involved with Claims Training. As I observed in my notes of the first meeting April 3, 1994, the session moved along slowly, crossing old ground several times with most of the members being cautious. This lack of clear focus led to several discussions involving emotional issues surrounding the need for change. For example, from my field notes of April 8, 1994, one member whose former group would be affected by changing the Claims Training Department was very willing to express his views and it was obvious that he had strong emotions surrounding the changes and the manner in which the changes would take place. This had a strong

influence on the committee and made the committee sensitive to the impact of change upon Claims Training Department trainers. While likely unintended, the focus of the evaluation stayed away from performance evaluation, which would have reflected on the individual trainers. The evaluation was therefore directed towards the product being delivered rather than the delivery of the product. In effect, the trainers were protected from criticism by the focus of the evaluation being on the product.

Conversely, the Litigation Management Evaluation Task Force Committee had a very clear mandate. The Litigation Management Evaluation Task Force Committee, apparently intuitively, was able to follow the Provus Discrepancy Evaluation Model precisely. Research by the Litigation Management Evaluation Task Force Committee at the outset led to a review of Claims Division Bulletin 894 which laid out objectives, procedures and expected benefits of the programme.

The Litigation Management Evaluation Task Force Committee's mandate was to determine compliance by B.I. (Bodily Injury) Adjusters and stakeholders to Bulletin CDB894 and its effectiveness. The Litigation Management Evaluation Task Force Committee spent the first meeting both in placing the evaluation in an historical perspective and in defining the problem to be examined. As my field notes

indicate, "It was not necessary to define the design criteria as the criteria were specified in Claims Division Bulletin 894."

Table 2

Comparison - Provus Discrepancy Evaluation Model, Stage I,
and the Litigation Management Evaluation Task Force

Committee Mandate

Personal Field Notes	Litigation Management Evaluation Task Force Mandate
<p>It was not necessary to define the design criteria as the criteria were specified in Bulletin CDB894.</p>	<p>The mandate of the Litigation Management Evaluation Task Force Committee was to reduce claims allocated expenses and severities by:</p> <ul style="list-style-type: none"> - clarifying and limiting involvement of external defense counsel in the settlement process. - arranging earlier potential end dates (discovery and trial) for the legal process, thus promoting earlier settlement. - implementing a process by which counsel and adjuster will agree to a documented and budgeted course of action for each new litigated file requiring more than closure of pleadings. - deferring the initial counsel review of complete file contents unless and until necessary. - providing further tools for defense counsel evaluation and audit.

Table 2 above outlines the Mandate as it was presented, by Senior Management, to the Litigation Management Evaluation Task Force Committee.

This well defined objective accords with the Provus Discrepancy Evaluation Model and lays the groundwork for Stage II, Programme Operation. After having defined an objective it is crucial that a similarly well defined implementation plan be designed along with an effective evaluation process which must be in place in order to establish accurately whether or not the results attained are congruent with the original objectives.

Stage II -Programme Operation. Stage II of the Provus Discrepancy Evaluation Model compares the current operations of the programme being evaluated with the original objectives and procedures set out in the design criteria. It is necessary to examine the results attained against the original objectives. According to Provus, only by establishing whether or not congruence exists between the original objectives and the attained results can an evaluation effectively establish the success or failure of a programme.

Stage II of the Provus Discrepancy Evaluation Model, Programme Operation, was not followed by the Claims Training Evaluation Task Force Committee. Instead of following steps similar to those suggested by Provus, which would have

required the conducting of a comparison between the intended objectives and the actual outcomes of Claims Training, the Claims Training Evaluation Task Force Committee chose to examine current Claims Training students and stakeholder groups through the use of focus groups and individual interviews to determine what the expectations were of Claims Training. As stated in my field notes of April 8, 1994:

Given the expense of maintaining a centralized programme and a demand from the field for training upon specific day-to-day issues, the question that surfaced was whether or not there was a better way to conduct staff training.

The Claims Training evaluation was predicated upon an assumption of a perceived problem, Claims Training's lack of adding value to the students' development, and sought to define alternate methods of delivery which included an emphasis upon centralized training versus decentralized training.

This movement directly into the investigation of a perceived problem and a movement away from the examination of actual versus intended outcomes left the Claims Training Evaluation Task Force Committee open to criticism with respect to the outcome of the evaluation.

On the other hand, the Litigation Management Evaluation Task Force Committee created a simple four-question template (Table 3, right hand column) which was designed to elicit answers from a selection of Bodily Injury Adjusters and

stakeholders interviewed which would establish compliance with Bulletin CDB894, the official policy bulletin which set out a litigation budgeting procedure.

This is similar to the suggestion in the Provus Discrepancy Evaluation Model which establishes a foundation for evaluations by identifying what the intended outcomes are and comparing them to the actual outcomes. The Litigation Management Evaluation Task Force Committee evaluated the effectiveness of Bulletin CDB894 through the design of a simple evaluation questionnaire to determine whether or not there was compliance.

The Litigation Management Evaluation Task Force Committee used a small sampling of staff members and stakeholders to answer the four questions, outlined in Table 3 below, to project the potential of whether or not there was general compliance on a corporate-wide basis.

The Litigation Management Evaluation Task Force Committee followed, apparently intuitively, the procedures suggested in Stage II of the Provus Discrepancy Evaluation Model wherein Programme Design was examined in terms of input and process dimensions.

The current operation at that time was compared with the design criteria by the Litigation Management Evaluation Task Force Committee in order to determine compliance.

Table 3

Comparison - Provus Discrepancy Evaluation Model, Stage II,
and the Litigation Management Evaluation Task Force

Committee Mandate

Personal Field Notes	Litigation Management Evaluation Task Force Committee Final Report Page 7
The current operation was compared to the objectives and procedures as laid out in CDB894.	<p>The Litigation Management Committee decided that the best method of obtaining an accurate assessment of the state of the current litigation management practices was to review CDB894 and interview BI Adjusters and Defence Counsel. The Committee would also conduct a review of the literature of the reference material available on the issue of litigation management. Defence Counsel and BI Adjusters were asked to respond to the following questions:</p> <ol style="list-style-type: none"> 1. Do you comply with CDB894? 2. What about CDB894 works or doesn't work? 3. Are the requirements of CDB894 necessary? If not, what are the alternatives? 4. What can be done to improve CDB894?

In other words, the Litigation Management Evaluation Task Force Committee identified what the intended outcomes were

by confirming the requirements of Bulletin CDB894 and compared them to the actual outcomes in order to establish whether or not a variance existed.

Stage III - Programme Interim Products. Stage III of the Provus Discrepancy Evaluation Model requires an examination of the process and the specific outcomes. The focus of the Claims Training Evaluation Task Force Committee turned from the outset towards conducting a needs assessment and not towards defining the process and intended outcomes of claims training.

As was evidenced by the methodology used by the Claims Training Evaluation Task Force Committee, the Claims Training Evaluation Task Force Committee appeared to make the assumption that claims training was dysfunctional and not meeting the needs of the recipients. One result of this assumption was that the questions put to the focus groups steered the answers towards what the Claims Training Evaluation Task Force Committee appeared to feel was right or wrong about the Claims Training Department. A second result was that the responses recorded by the Committee, when interpreted post interview, appeared to have been influenced by those assumptions.

The following examples of questions used illustrate the apparent bias in the evaluation design.

- Is this specific to the BIT (Bodily Injury Training) Programme or is it training as well as BI's in general. [Questions whether the perceived problem rests with training or the adjusters themselves] Is this a revision of the BIT programme?
- Organization structure- does it matter?
- Would it be different if learning centre was out of office?
- What claims training is needed?
- Training in advance of need - is this a problem?
- Should there be a new, specialized course?
- Would there be some value to a networking set-up?
- Decentralizing or regionalizing - no travel time - do you have any comments on this?
- What if, instead of going to HO you had a regional office - would that provide the same thing?
- Regionalization-should training be at office site, or off site within region?
- Is what they teach worthwhile?
- Do trainees come out of the training programme with the basic knowledge/theory they need?

The types of questions asked were not those which would compare the outcome of the results of claims training with the intended outcomes. Rather, these questions took the respondents into a direction which would redefine claims

training. While these questions were valid from the perspective of determining a new focus for claims training they were premature and were the type of questions which could be asked after establishing whether or not claims training had met its original objectives. In other words, these types of questions were more appropriate for designing a new direction of claims training after determining that a change of focus was required.

The Provus Discrepancy Evaluation Model requires that actual outcomes be compared to intended outcomes which would focus an evaluation in an entirely different direction from that taken by the Claims Training Evaluation Task Force Committee. This stage of the Provus Discrepancy Evaluation Model underscores the need for developing clearly defined objectives.

Ralph Tyler in his article Changing Concepts of Educational Evaluation (1986, pp. 53-55) identifies four criteria for developing objectives. They are summarized as follows:

1. Objectives should be in harmony with the educational philosophy of the school.
2. Relevance and appropriateness of the objective to the subject matter.
3. The opportunity the learner has to use what he or she is learning.

4. The appropriateness of the objective to the needs, interests and present development of the particular students for whom the program is planned.

Tyler (1986, p.55) makes the following comment with respect to the importance of clearly defining objectives.

The process of evaluating the objectives of a proposed educational program is largely that of reminding those responsible for the development of the program that these four criteria should be carefully considered.

The results of the focus group sessions formed the basis of the recommendations of the Claims Training Evaluation Task Force Committee. For example, from the May 6, 1994 focus group session the following sample of quotations made by the respondents also illustrate the focus of the evaluation.

- We could have easier access to resources.
- Biggest concern re initial training -- too far away from actual work we do.
- Hiring someone from the outside who doesn't know claims process, their training is not specific to the problem at hand.
- Lots of "war stories".
- Learning through PC - no.
- When you are at HO, they don't know what is going on in claims needs to be more of a "together" working relationship.

- Claims training - very entertaining! Perhaps we needed more direction.

In contrast to the procedures used by the Claims Training Evaluation Task Force Committee, the Litigation Management Evaluation Task Force Committee interviews, which were conducted with B.I. (Bodily Injury) Adjusters and stakeholders, were specifically directed towards determining whether a variance existed between the intended outcomes and the actual outcomes of the application of Bulletin CDB894. The questions developed in Stage II were used and the interviewees were allowed to elaborate upon their answers. As the Litigation Management Evaluation Task Force Report indicates on page 4, "it rapidly became apparent that there was virtually no compliance" with Bulletin CDB894 and, in general, that the working relationship between Defence Counsel and B.I. (Bodily Injury) Adjusters was strained. These results were presented in the final report to senior management and recommendations were made to terminate Bulletin CDB894 with the intention of implementing a new litigation management process.

Stage IV - Programme Terminal Products. In Stage III of the Provus Discrepancy Evaluation Model, Programme Interim Products, the Claims Training Evaluation Task Force Committee focussed upon needs assessments rather than defining the process and defining the intended outcomes of

claims training as is required by the Provus Discrepancy Evaluation Model. The Claims Training Evaluation Task Force Committee began evaluating the Claims Training programme output design against the expectations of the students and stakeholders. While this reflects Stage IV of the Provus Discrepancy Evaluation Model and would come towards the end of the process, the Claims Training Evaluation Task Force Committee moved into this stage almost immediately. The first interviews took place on the April 20, 1994, 12 days after the formation of the Task Force. The first focus group session was held on May 6, 1994, 28 days after the formation of the Claims Training Evaluation Task Force Committee.

The following excerpt from the Focus Group Facilitator's Guide (April 6, 1994) crystalizes the direction in which the Claims Training Evaluation Task Force Committee chose to take this evaluation.

The facilitator may want to use probing questions associated with the three objectives to clarify input.

What Claims training is needed? Probes: What would be different from the way it is now? For whom (CR's [Claims Representatives], CA's [Claims Adjusters], BI's [Bodily Injury Adjusters])? What would add, delete, or change concerning content, amount, or timing (appropriate timing for the various stages of BI or CA development)? WHY? Is the current training relevant to the job, effective/not effective, up-to-date? (Any comments around learning methods: lecture, interactive, case study, on-the-job, etc.?)

What is the best organization structure to provide the training? Probes: which division or department would it be best to have own the claims training (eg. HR-hrd [Human Resources-Human Resources Development]; Claims-Field, head office, MD R&T [Material Damage-Research & Training])? WHY?

Who should deliver what claims training? Probes: HRD Trainers, UM's [Unit Managers], external consultants/trainers, distance education, PC based, degree or type that could be centralized/decentralized? WHY?

The following excerpts are typical of the responses to the questions which were put to the focus groups by the Claims Training Evaluation Task Force Committee. They further illustrate the direction in which the focus groups were headed and highlight the lack of focus on a clear mandate.

May 6, 1994-BI Focus Group:

- Claims Training is too far removed from the field.
- Current system was good for initial training but there is a need to provide training/assistance on real life files in day to day operation. "XX", to the best of his schedule, helped in this area.
- We've been reactive and a proactive training approach is necessary.
- Wednesday mornings set aside for training.
- Easier access to resources required.
- User friendly access to information.

The diversity of the responses in these sessions, as outlined above, is an indicator of the lack of clear focus for the evaluation with respect to whether or not the Claims Training Department had achieved its objectives.

The direction that the Claims Training Evaluation Task Force Committee took with the data gathering process was based upon assumptions that were made about the successes or failures of the Claims Training Department. The format of the focus group sessions was aimed at examining processes and establishing whether or not the Claims Training Department was meeting the current needs of the staff. To this end, the focus group process was designed, intentionally or not, to fulfil the assumptions made by the Claims Training Evaluation Task Force Committee. As the above answers show, the results appear to have been gathered through the use of leading questions to achieve specific outcomes.

The Claims Training Evaluation Task Force Committee focussed upon the fulfilment of needs. The Provus Discrepancy Evaluation Model, however, focussed upon identifying the objectives (intended outcomes). There is a fundamental operational difference in approach between the Claims Training Evaluation Task Force Committee and the Provus Discrepancy Evaluation Model. The Claims Training Evaluation Task Force Committee took the fulfilment of

current needs as being the measure of success. Needs which exist at the time of the evaluation, while potentially interesting, are irrelevant in determining whether or not the programme objectives achieved what they set out to do. The emphasis should be, as the Provus Discrepancy Evaluation Model dictates, upon programme objectives and whether or not they were achieved.

The Claims training evaluation moved towards conducting a needs assessment, establishing a benchmark of the then current ability of the Claims Training Department to service its customers adequately and created a vision of what the future might look like. However, it did not evaluate the Claims Training Department's success at achieving its objectives as those objectives were never discussed nor explored.

The Litigation Management Evaluation Task Force Committee had no difficulty in following through in its apparently intuitive mirroring of the Provus Discrepancy Evaluation Model. The questions that the Litigation Management Evaluation Task Force Committee developed were very specific and were clearly focussed upon determining whether or not there was compliance with Bulletin CDB894. There was no obvious bias in the questions and the results were simple to interpret as they essentially only required a "yes" or "no" type answer.

There were major differences between the manner in which both committees proceeded with their evaluations. These differences may have developed due to the differences in their mandates as well as the membership of each committee. In other words, different mandates may require different approaches. The members of the Claims Training Evaluation Task Force Committee were made up of a cross section of staff from the corporation and were led by a chairman with Head Office experience. The Litigation Management Evaluation Task Force Committee consisted of members of the corporation from a very specialized area with field experience. The differing mandates and perspectives could have influenced the development of the two different approaches taken.

Stage V - Programme Cost. No cost analysis was made by the Claims Training Evaluation Task Force Committee but recommendations were advanced to examine the cost benefits of maintaining permanent in-house training versus "just in time" training which could be provided by contracting with outside consultants. This recommendation was developed during the Claims Training Evaluation Task Force Committee's wrap up session on May 31, 1994. It, among other recommendations, was formulated through a synthesis of trends and observations. This recommendation was put

forward as a serious suggestion that would require further development.

While the Claims Training Evaluation Task Force Committee recommended, apparently as an afterthought, that a cost analysis be conducted, this is actually a planned part of the process as set out by Provus. It is not a process to be referred to and passed along to some future committee or evaluation team but is rather an integral part of the evaluation process. An evaluation using the Provus Discrepancy Evaluation Model is not complete until the cost analysis has been conducted.

While a cost/benefit analysis was not performed by the Litigation Management Evaluation Task Force Committee, recommendations were made that one be completed in order to "measure the effectiveness of procedural change." These recommendations were made almost as an afterthought during the composition of the final report. There was no evidence in the committee minutes or my notes that this was an earlier consideration. As the author of the final report, I added the cost/benefit analysis recommendation to the Litigation Management Evaluation Task Force Committee report. I included the cost/benefit recommendations as I was aware that this was a necessary part of an evaluation from previous training.

Stufflebeam et al. (1973) C.I.P.P. Evaluation Model

The Provus Discrepancy Evaluation Model serves to illustrate the importance of defining clear objectives in order to focus an evaluation. Similarly, the Stufflebeam et al. C.I.P.P. Evaluation Model establishes evaluation guidelines which underscore the need to have a clearly focussed evaluation by placing the evaluation in context and examining the input, processes and product of the evaluation (C.I.P.P.). Both the Claims Training Evaluation Task Force Committee and the Litigation Management Evaluation Task Force Committee were examined from the perspective of the Stufflebeam et al. Model to determine if parallels exist between theory and practice.

Context Evaluation. Stufflebeam et al., in the C.I.P.P. Evaluation Model, take the position that the first stage of the evaluation should be to place the object being evaluated in context. This involves identifying and assessing needs within the context of the evaluation and defining problems which underscore the needs.

The Claims Training Evaluation Task Force Committee did conduct an unstructured needs assessment through the focus group process, from May 06, 1994 to May 20, 1994 but did not place it in context as the actual programme objects were not examined. The needs assessment was based upon examining the responses made during the focus group process to identify

those responses that related directly to the needs of the respondents. The process undertaken by the Claims Training Evaluation Task Force Committee to determine needs could not be considered a formal needs assessment but it does indicate that a needs assessment was considered. There were leading questions to guide the focus groups but the format of the sessions was free flowing in order to establish a comfortable atmosphere and encourage creative thought.

On the other hand, the Litigation Management Evaluation Task Force Committee did identify the standards expected by examining Bulletin CDB894 and placed them within an operational context. As Bulletin CDB894 was a published document, it saved the Litigation Management Evaluation Task Force Committee a considerable amount of time in that it was very clear what was to be evaluated and what the expected results would be.

Input Evaluation. The second stage, Input Evaluation, identifies and assesses system capabilities, available input strategies, and designs for implementing the strategies. In simple terms, the system in which the object being evaluated and is expected to operate is examined to determine if the existing infrastructure is capable of supporting the programme being evaluated. The infrastructure includes logistical issues dealing with physical needs as well as supporting programmes and operational issues which consist

of programme design and implementation strategies. If these are not available, then there is little point in continuing the evaluation. If the system is incapable of allowing for implementation to take place, then according to Stufflebeam et al., the evaluation would cease at this point.

The Claims Training Evaluation Task Force Committee did assess the availability of resources and identified implementation strategies by recounting the then-existing structure and methods of delivery of the Claims Training Department. This was accomplished at the outset of the Claims Training Evaluation Task Force Committee on May 5, 1994. As an example, in an interview on May 5, 1994 with a senior manager, the Claims Training Evaluation Task Force Committee captured the following comments.

-We are in a time of restraint - but don't let staffing issues influence where your recommendations are going. If we need more staff, we will take it under advisement. Don't think that just because it requires additional resources, you shouldn't recommend. Up to us to find resources to do it.

-To have two departments doing technical training in isolation doesn't help the overall effort. Should be a closer link. Needs more integration from training perspective.

These recommendations were process oriented and did not reflect an evaluation of the Claims Training Department of that time, but rather sought to fill in gaps between staff needs and delivery. Three examples of the nine

recommendations from page 10 of the Claims Training Report illustrate the focus on redesign as opposed to evaluation.

1. That Claims Education Services become part of the Claims Division as part of a new department titled "Claims Training and Research".
2. That the Manager, Claims Training and Research, report directly to the Vice President, Claims. This Department would be responsible for integrating and meeting all training needs within the Claims Division.
9. That each Claim Centre have a Claims Office Training Liaison person, whose responsibilities will include the co-ordination of the training needs of all Work Groups within the Claim Centre. p. 10

The Stufflebeam et al. C.I.P.P. Model is concerned with evaluation, not the re-design or re-engineering of the object being evaluated. This leads to a fundamental difference between the Claims Training Evaluation Task Force Committee approach and Stufflebeam et al. The Stufflebeam et al. C.I.P.P. Evaluation Model takes a clinical approach and seeks to diagnose by establishing whether or not the intended outcomes were achieved within the context that the programme being evaluated operated. On the other hand, the

Claims Training Evaluation Task Force Committee assumed that a problem existed (i.e., training not meeting needs) and sought to cure it by developing recommendations for change which would result in a "new" training department being designed. This was complete all the way from what should be taught, to whom it should be taught right up to the reporting structure of the newly designed department. In the end, the Claims Training Evaluation Task Force Committee never did address whether or not the Claims Training Department had met its objectives.

The Litigation Management Evaluation Task Force Committee also conducted a loose system analysis. As all members of the Litigation Management Evaluation Task Force Committee had worked for the corporation and functioned within its corporate structure for several years they were intimately familiar with the then existing corporate structure. Whether or not the structure was capable of supporting the implementation of Bulletin CDB894 was never discussed. This absence of consideration would appear to run contrary to the Stufflebeam et al. C.I.P.P. Evaluation Model. It may be, however, that the closeness of the evaluators to the situation being evaluated could allow for some unspoken "givens" which are implicit and may not be apparent to a neutral observer from outside the Claims Division nor possible had the evaluation been conducted by

an external evaluation team. This aspect of advantages/disadvantages of internal or external evaluators is not considered by either the Provus or Stufflebeam et al. Evaluation Models. This is interesting because if the evaluation had been conducted by evaluators who did not have an intimate understanding of the structure they would have had to complete a detailed analysis. Being familiar is advantageous in that it saves time but important issues could be overlooked. The Stufflebeam et al. model, used in a strict application, does not take into consideration external factors which may have an impact upon the supporting infrastructure. Therefore, while the structure may exist to support implementation, implementation does not take place. The capability is there, but circumstance does not allow it to happen. In the case of the Litigation Management Evaluation Task Force Committee, for example, while the structure was sufficient to support the initiative, there were external factors which competed for the time of the people expected to implement the programme. Consequently, non-compliance was encouraged due to a natural prioritizing of functions. If the programme was not given a high profile to emphasize its high priority, little or no compliance could be expected. On the surface, therefore, the structure appeared adequate to those close to the structure, but time and position priority may have been

contributing factors to non-compliance. The Litigation Management Evaluation Task Force Committee determined that while the system was capable of supporting the implementation of Bulletin CDB894, it was the consensus of the Task Force that the implementation procedures were cumbersome and would encourage non-compliance. Below are three examples of this prediction which come from the Litigation Management Evaluation Task Force Committee meeting minutes of October 19, 1993:

- Ramifications of going over budget not identified/ explained to defence counsel/adjusters
- Accountability without authority
- Poor implementation or follow up

Very little time was spent in this area due to the familiarity of the Task Force with the corporation's field staff capabilities and support infrastructure. In fact, there was only a passing mention of this "analysis". As an e-mail to a committee member from a staff member who was surveyed illustrates:

in theory it is a great idea, but in practice didn't work. perhaps this is because the corp. let the thing drop that it never got past the "working out the bug stage". maybe if we insisted it be done on all files, and updated as the file progressed, it may have developed into a good planning and cost control tool. (E-mail, 19th of October, 1993)

In essence, as required by the Input Evaluation stage of Stufflebeam et al. C.I.P.P. Evaluation Model, the system was

capable of the implementation of Bulletin CDB894 and sustaining it, but, as Stufflebeam et al. fail to take into account, the lack of commitment, follow-up and the potential influence of external factors led to non-compliance. This failure to take commitment, followup and the influence of external factors into consideration is a limitation of the Stufflebeam et al C.I.P.P. Evaluation Model.

Process Evaluation. In the Stufflebeam et al. C.I.P.P. Evaluation Model, Process Evaluation involves the prediction of defects in the procedural design or implementation strategies and maintains a record of procedural events and activities.

The Claims Training Evaluation Task Force Committee made no prediction of programme design flaws. This absence of prediction is a reflection of the single-minded direction that the Claims Training Evaluation Task Force Committee took in assuming that a needs assessment would illuminate what was or was not working with Claims Training. The meeting minutes and my field notes of April 8, 1994 reflect that there was no evidence to support that a prediction of programme design flaws was considered. No monitoring of potential procedural barriers took place; therefore, the Claims Training Evaluation Task Force Committee was blind to alternate courses for the outcome, positive or negative, of Claims Training.

On the other hand, the Litigation Management Evaluation Task Force Committee did make predictions as to the success of compliance with Bulletin CDB894 and stated a probable cause. The Litigation Management Evaluation Task Force Committee meeting minutes of October 18, 1993 indicate that the discussions centred upon why Bulletin CDB894 failed and listed the following direct quote of the predictions of why Bulletin CDB894 failed in the October 18, 1993 committee minutes.

- drew battle lines between adjusters and defence counsel
- Defence counsel felt too much emphasis on budget \$\$\$'s
- Defence couysel felt they would be locked into budget when unable to predict variables.
- Ramifications of going over Budget not identified/explained to Defence Counsel/ Adjusters
- Conflict of message
- Conflict as to when Budget/Planning process to take place, i.e., before or after Examinations for Discovery.
- Accountability without Authority
- Bulletin being ignored
- Unrealistic time frames
- In-house Counsel exempt
- No rationale for issuing bulletin
- Defence Counsel (external) felt there was a hidden agenda
- Alienated Defence Council from Adjusters
- Process take focus away from file resolution
- Poor implementation and/or follow-up
- No consideration for other programs

[Note:- the above typographical errors are as they appear in the original document]

The Litigation Management Evaluation Task Force Committee felt that there would be little if any compliance

with Bulletin CDB894. As stated above, there was no rationale for issuing the bulletin and with conflicting messages the procedures were just too cumbersome to be effective. This prediction of non-compliance arose from a roundtable discussion on October 18, 1993 of the effectiveness of Bulletin CDB894, and relied heavily upon the experience of those committee members who had been involved with the corporation's success rate of similarly complex programmes. In the past, programmes had been developed by head office departments for implementation in the field. As well constructed as they were, they often failed to have a field perspective or consideration of field priorities. Given lack of follow-up, buy-in and ownership from the field staff, these head office programmes often received a lower priority from the staff whose time and demands were driven by servicing the public and ensuring good statistical results. The head office programmes had to compete with the day to day business demands with which the field staff have to contend. In other words, the expectation that staff and defence counsel would comply with the implementation of a budgeting process as outlined in Bulletin CDB894 was compared to what was actually occurring in the field. The conclusion was that Bulletin CDB894 had failed, and a number of reasons for failure, in addition to those predicted, were identified and are referenced on page

five of the Litigation Management Evaluation Task Force Committee's final report and are listed in Table 4 below.

Interestingly, these complaints revolve around general working relationships rather than Bulletin CDB894. The Stufflebeam et al. C.I.P.P. Evaluation Model assumes that an adequate infra-structure must be in place, but fails to consider other factors as indicated above. In this case, the structure was fine but other factors prevented implementation.

TABLE 4

Bulletin CDB894 Failure: Reasons For Failure

Defence Counsel	Bodily Injury Adjusters
<ul style="list-style-type: none"> - Adjuster failing to provide adequate initial instructions on assignment of files to defence counsel. - The constant changing of adjusters assigned to files. - Adjusters making unreasonable demands upon defence counsel; chambers (Court) applications, - Lack of communication with B.I. Adjusters (Bodily Injury Adjusters) 	<ul style="list-style-type: none"> - Defence counsel failing to follow instructions. - Defence counsel not focusing upon early settlement opportunities. - resistance from defence counsel to aggressively manage litigated files. - Junior counsel being assigned to files. - Changing counsel during the course of litigation. - Having to pay defence counsel to review the file material after a new counsel has been assigned due to changes within the defence firm. - lack of communication with defence counsel.

This follow-up session relied upon personal anecdotes as well as the staff and defence counsel interviews. The Litigation Management Evaluation Task Force Committee prediction and the specific interview questions which were put to those interviewed led to a very focussed evaluation which resulted in savings in both time and effort. In essence, the Litigation Management Evaluation Task Force Committee clearly defined what needed to be evaluated, predicted what the outcome would be and then implemented an evaluation procedure to test the prediction. This focus prevented a wasting of time and effort by all involved with the evaluation.

Product Evaluation. The Claims Training Evaluation Task Force Committee could make no comparison between activities recorded and actual programme objectives because the actual programme objectives were never examined or defined. Rather, recommendations were made based upon the assumption that an examination of the personal needs of the stakeholders and recipients of Claims Training would constitute an evaluation of the Claims Training Department. The only measurement that took place involved a comparison between outcomes and expectations. The recommendations that were made were based solely upon this comparison. Based upon the assumption that the current Claims Training Department was not effective due to the lack of satisfaction

of needs, a new structure was proposed. While the recommendations may or may not have been beneficial, the evaluation itself is open to procedural question. The recommendations were made with the intention that future committees would take over ownership of the problem with respect to implementation.

The Litigation Management Evaluation Task Force Committee went beyond making this single recommendation to recommend that the whole litigation management process be re-engineered. The Litigation Management Evaluation Task Force Committee, while recognizing that it could not take on the full responsibility for a complete overhaul, did create sub-committees to look at specific aspects of the litigation management processes that had been identified as being high priority issues. The Litigation Management Evaluation Task Force Committee, in the fall of 1994, debated the scope of the Litigation Management Evaluation Task Force Committee mandate. After some debate, the Litigation Management Evaluation Task Force Committee decided to create sub-committees to explore issues that surfaced during the evaluation which were important enough to be addressed but which were not directly tied to the Litigation Management Evaluation Task Force Committee's mandate. The Litigation Management Evaluation Task Force Committee final report (page 8) made recommendations that specialized defence plans

be created as an offshoot of the Litigation Management Evaluation Task Force Committee mandate.

During the course of the committee's investigation into litigation file management, the committee was asked to develop "canned defence plans". As a result, a sub-committee was struck and a defence plan for fibromyalgia (a condition which is being associated with traumatic injury and receiving high dollar awards in court) is being developed.

The Litigation Management Evaluation Task Force Committee took on the responsibility of re-engineering the litigation file handling process, developing both new and simpler forms and procedures. The changes were implemented in a test situation and formally adopted corporate-wide. Regular follow-up was conducted to make modifications where required to ensure simplicity and compliance.

Neither Provus nor Stufflebeam take into account the external influences or personal agendas which may impact an evaluation. By not addressing these issues both models are excluding the potential impact that external influences may have upon the application of their evaluation models.

Summary

The results presented in Chapter Four illustrate the effectiveness of applying two theoretical evaluation models to two evaluations conducted in a corporate setting. The Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model were examined, stage by stage, and compared to the operational stages of both the Claims

Training Evaluation Task Force Committee and the Litigation Management Evaluation Task Force Committee evaluations.

This comparison revealed that the Claims Training Evaluation Task Force Committee did not follow the pattern of either theoretical model. The Litigation Management Evaluation Task Force Committee's evaluation was very similar to those of Provus and Stufflebeam et al. This study also identified gaps in the effectiveness of both the educational evaluation models that could have had an impact on the outcome of the evaluations had they been actually applied.

Chapter Five will discuss the results of this study and consider the benefits of using formal educational evaluation models. In particular, the benefits of using the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model will be examined within the context of this study. Differences and similarities between the two theoretical educational evaluation models and the two corporate based evaluations will also be explored.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

In Chapter Four, the Claims Training Evaluation Task Force Committee and the Litigation Management Evaluation Task Force Committee evaluations were assessed within the frameworks provided by the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model. There were differences and similarities between the theoretical evaluation models and the manner in which these two corporate based evaluations were applied. Chapter Five will summarize the results of this study, discuss the relevance of these findings in the conclusion and make recommendations for further study. Two thoughts to consider when reading Chapter Five are:

1. How the two evaluation committees could have benefited from using an evaluation model such as either the Provus Discrepancy Evaluation Model or the Stufflebeam et al. C.I.P.P. Evaluation Model.
2. There are concerns with respect to what appears to be deficiencies in the practical application of both Provus's and Stufflebeam et al.'s theoretical models. Two perceived weaknesses that have been identified are "Stall Points" and the "Impact of External Forces on Evaluation Model Effectiveness."

Summary

Identification of Programme Objectives and Placing Them in Context. The Provus Discrepancy Evaluation Model requires that the objectives of the programme being evaluated be clearly defined. These programme objectives must be identified prior to the evaluation proceeding in order that the evaluators have no question in their minds as to what it is that they are evaluating. This is necessary to provide a proper focus on the evaluation. If the objectives are unclear or misinterpreted, the evaluation will be misdirected and the value of the results of the evaluation would be questionable. By properly applying the Provus Discrepancy Evaluation Model an evaluation team would be able to concentrate upon determining whether or not the programme being evaluated had achieved what it set out to accomplish.

For example, the Claims Training Evaluation Task Force Committee did not spend time specifically identifying and clarifying the objectives of claims training. This led the committee to make assumptions about claims training and conduct its evaluation in such a manner as to support or reject those assumptions. Given that the assumptions made may or may not reflect the objectives of claims training,

the conclusions reached by the evaluation are weak at best and highly subject to criticism.

The Litigation Management Evaluation Task Force Committee, on the other hand, was provided with specific written objectives with respect to determining compliance with Bulletin CDB894. The objectives were very clear and the evaluation was focussed upon whether or not there was compliance. The results of the evaluation were compared to the original objectives and there can be a measure of confidence with the conclusions. The interesting point to note with the Litigation Management Evaluation Task Force Committee is that no formal evaluation model was applied and it appears that this committee, by circumstance, "fell into" a process that was similar to a formal evaluation model.

The Stufflebeam et al. C.I.P.P. Evaluation Model places the evaluation in an operational context and assesses needs within that context. By placing the evaluation in context, it is possible to compare actual outcomes with intended outcomes. An evaluation team utilizing the Stufflebeam et al. "Context Evaluation" stage will not only be able to identify clearly the programme objectives as suggested by Provus, but will also be able to place the design objectives within a contextual framework which would delineate the objectives from the outcomes. This will allow the evaluation team to determine not only what the objectives

were, but whether they are synchronous with the intent of the programme.

For example, the Claims Training Evaluation Task Force Committee did not clearly identify the objectives of the Claims Training Department nor did it place the programme in the context within which it had to operate. There was no examination of the intent of the programme. This lack of clear objectives and context led to what appeared to be erroneous assumptions being made about the nature of the committee's mandate and subsequently, a misdirection of the evaluation.

The Litigation Management Evaluation Task Force Committee, on the other hand, was provided with clear objectives. The committee members were aware that compliance or non-compliance with Bulletin CDB894 had financial implications as Bulletin CDB894 was a corporate directive that set out a procedure for the litigation management process. Again, the Litigation Management Evaluation Task Force Committee, apparently intuitively, followed an evaluation process without applying a formal evaluation model.

Examination of Programme Structure and Supporting Infrastructure. By examining the structure of the programme being evaluated in terms of its operational context in comparison to the programme design criteria, it

is possible to determine whether or not the programme infrastructure is capable of supporting the programme. Stufflebeam et al. assist the evaluation team by highlighting this as a "Process Evaluation" stage. The benefits of clearly defining the programme implementation structure are twofold. Firstly, as Stufflebeam et al. suggest, if the supporting infrastructure is found to be logistically inadequate, the evaluation team need not proceed further. This conclusion could potentially lead to the infrastructure being modified to ensure adequate resources in order to allow the programme to proceed and enhance the possibility of the objectives having an opportunity to be developed.

Secondly, as Provus indicates, the programme operation is compared to the standard of the programme design to ensure that there is congruence with the design criteria. In this case, the supporting infrastructure may be logistically adequate, but if the design of the supporting infrastructure is such that it causes the programme to stray from the intent of the objectives then once again the evaluation should cease at this point. The programme should be terminated or at the very least, the infrastructure should be modified to allow the design criteria to be supported.

Anticipation of Barriers to Success. As Stufflebeam

et al. purport, it is an important feature of evaluation to examine and anticipate what may be barriers to success and where they may lie within the evaluation process. The benefit here for an evaluation team is that without limiting the scope of the evaluation it directs the evaluators towards likely problem areas that, if confirmed, would allow the evaluation to be terminated. In other words, the whole evaluation process could be shortened. For example, the Claims Training Evaluation Task Force Committee did not make an attempt to identify potential barriers and as a result, their efforts were spread out in a number of directions. The end result was a very broad approach to the evaluation in order to determine whether or not the programme was successful.

On the other hand, the Litigation Management Evaluation Task Force Committee did anticipate barriers to the programme being evaluated and this led the evaluation team to focus its attention upon the perceived barrier and either prove or disprove its theory. The theory, in this case, was proved and the whole evaluation process was shortened. The Litigation Management Evaluation Task Force Committee was able to make more efficient use of time, resources and money by shortening the evaluation process through the

anticipation and verification of potential barriers to the success of the programme being evaluated.

Orientation of Programme Towards Results. Both the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model are results oriented. Provus examines the results in relation to the stated programme objectives. Results were important to both the Claims Training Evaluation Task Force and Litigation Management Evaluation Task Force Committees. The results of the Claims Training Evaluation Task Force Committee were the product of an evaluation based upon assumptions of questionable value and are therefore themselves questionable. The Litigation Management Evaluation Task Force Committee results could be compared with specific objectives and would add value to the evaluation. The process, once again, was due more to luck and circumstance than to design.

Similarly, Stufflebeam et al. examine the results from a contextual perspective. In other words, does the programme achieve what it states that it set out to accomplish? This reflects upon the need to have clearly defined objectives to determine the success of the results.

Conclusions

Benefits of using Provus and Stufflebeam et al. Evaluation Models. Both committees could have enjoyed

definite benefits had they used either the Provus or Stufflebeam et al. Evaluation Models. Both committees would have realized benefits simply by the application of any evaluation model, regardless of its type as both committees lacked a formal evaluation structure and seemingly proceeded through their evaluations by trial and error. The Claims Training Evaluation Task Force Committee, in particular, strayed from formal evaluation principles and the basis for its conclusion is weak, which brings the results into question. The Litigation Management Evaluation Task Force Committee, apparently intuitively, followed a process that was similar to formal evaluation and its conclusions are seemingly well founded and solid.

While the Litigation Management Evaluation Task Force Committee did not use a formal evaluation process, and seemed to arrive at a similar outcome to what could have been expected had a formal process been used, it was more likely due to luck and circumstance than due to design. A formal evaluation process would have provided a solid framework that would have minimized the Litigation Management Evaluation Task Force committee's reliance upon luck.

Some of the benefits that the Claims Training Evaluation Task Force Committee and the Litigation Management Evaluation Task Force Committee could have

realized from utilizing the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model are: identification of programme objectives and placing them in context, examination of programme structure and supporting infrastructure, anticipation of barriers to success, orientation of programme towards results and financial comparison for cost containment.

Neither the Claims Training Evaluation Task Force Committee nor the Litigation Management Evaluation Task Force Committee formally examined the supporting infrastructure as suggested by Provus and Stufflebeam et al. This important stage of the formal evaluation process was given only a passing consideration by the Litigation Management Evaluation Task Force Committee (presumably because there was an inherent familiarity between the committee members and the infrastructure within which they worked on a daily basis) and the Claims Training Evaluation Task Force Committee did not consider it at all.

Both committees would have benefited from using formal evaluation processes, such as the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model, which would have examined this stage in detail. Had this stage been formally examined it may have revealed whether or not there were other factors which were

not considered that could have inhibited the effectiveness of either programme.

For evaluation teams using either model, the importance of examining the results in relation to the objectives is underscored by the outcomes of both the Claims Training Evaluation Task Force Committee and the Litigation Management Evaluation Task Force Committee. The Claims Training Evaluation Task Force Committee did not clearly identify the programme objectives nor did it reference its conclusions with objectives and the results were recommendations which have questionable value in relation to the success or failure of the programme that was evaluated. Future corporate plans may rely upon those recommendations that could have potentially negative results.

The Litigation Management Evaluation Task Force Committee, on the other hand, had a clearly defined objective and compared the results directly against that objective. Its conclusion with respect to the success or failure of the programme it evaluated was therefore better grounded than that of the Claims Training Evaluation Task Force Committee.

It is evident that for an evaluation to be effective the importance of having clearly defined objectives, which is identified by Provus and Stufflebeam et al. as a crucial element of a successful evaluation, must be stressed. The

comparison of the results achieved with the stated objectives is also a critical component in determining the accuracy of the evaluation's conclusions.

Financial Comparison For Cost Containment. Both Provus and Stufflebeam et al. have, as final elements of their models, a component dealing with cost analysis. They are relatively similar with respect to comparing the cost of the programme being evaluated with industry standards or similar programmes. This process may be termed "benchmarking". It is important in that regardless of the origins of the programme being evaluated, business decisions have to be made with respect to continuing or terminating programmes. In many cases, the basis for continuing or terminating a programme is financial.

Both the Claims Training Evaluation Task Force Committee and the Litigation Management Evaluation Task Force Committee considered but did not conduct a financial analysis. Both committees recommended that a financial analysis be conducted. The problem here, though, is that if a future financial analysis were conducted, it would likely not involve the original committee members and there would be a loss of continuity.

Other potential problems with both committees are a lack of completeness to the evaluation and a lack of

accountability for the implementation of recommendations that are made without a financial analysis.

For evaluation teams, there is a real benefit that may be had by following either the Provus or the Stufflebeam et al. Evaluation Models with respect to cost analysis. The models emphasize the importance of financial concerns that are addressed at the conclusion of the evaluation process. The placing of a financial assessment at the end of an evaluation illustrates that regardless of the success or failure of a programme through a comparison of objectives and results, the decision to continue or terminate a programme may be a financial one. It is the evaluation team's responsibility, unless otherwise directed, to complete the evaluation by conducting a financial analysis. The evaluation team is in a good position to conduct such an evaluation or give input to a financial evaluation committee after having conducted an in-depth programme evaluation. Failing to conduct a financial analysis, as was the case with both the Claims Training Evaluation Task Force Committee and the Litigation Management Evaluation Task Force Committee, leaves an evaluation incomplete.

Committees' Efforts Facilitated by using a Formal Evaluation Model

Some of the ways in which the Claims Training Evaluation Task Force Committee and the Litigation

Management Evaluation Task Force Committee could have facilitated their efforts by using a formal evaluation model are that a formal evaluation model provides structure, suggests methodology, helps with determination of results and aids with the making of recommendations.

Provides Structure. Both committees could have facilitated the operation of their committees and increased efficiencies by using a formal evaluation model such as the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model. The first benefit that is evident is that a formal evaluation model provides a structure and direction for an evaluation. By mapping out a plan for how an evaluation will take place, the evaluators would be better able to focus upon the crucial elements of the evaluation without distraction, misplaced resources and wasted effort. Wandering through an evaluation by designing it "on the go" is inefficient, risky and leaves both the process and the results open to criticism and challenge.

For example, neither the Claims Training Evaluation Task Force Committee nor the Litigation Management Evaluation Task Force Committee used a formal evaluation model. Had a formal evaluation model such as the Provus Discrepancy Evaluation Model or the Stufflebeam et al. C.I.P.P. Evaluation Model been used, it would have been apparent at the outset that the objectives of the programme

must be clearly identified. Failing to do this led the Claims Training Evaluation Task Force Committee to stray from evaluating the programme towards evaluating the suitability of the product of the programme, and making recommendations for product change. The committee did not address whether the programme, as designed, was successful in accomplishing what it intended to do given the original objectives. Using a formal evaluation model could have prevented this situation from arising and provided a focus to the evaluation.

In a like fashion, the Litigation Management Evaluation Task Force Committee could have had a similar benefit from using a formal evaluation model. While the Litigation Management Evaluation Task Force Committee seemingly followed the general outline of a formal evaluation, it did so on what appears to have been an intuitive level as opposed to a conscious effort. There was never a mention of following a specific evaluation plan or even setting out a formal process prior to the evaluation commencing. The Litigation Management Evaluation Task Force Committee seems to have been "lucky". The mandate of the committee was very simple and specific with the objectives of the programme being published as Bulletin CDB894.

The Litigation Management Evaluation Task Force Committee may have had similar difficulties to the Claims

Training Evaluation Task Force Committee if the objectives had not been specified in Bulletin CDB894. A formal evaluation model would have guided the evaluation process in an organized fashion without relying upon intuition as the rule.

Methodology. This study suggests that the results of an evaluation can be brought into question and potentially rendered of no value if the methodology of the evaluation can be shown to be flawed. In other words, evidence of the existence of an acceptable evaluation methodology is a crucial element of the evaluation process. The more sound and accepted the methodology is, the more confidence that may be given to the results. It must be remembered that the second part of methodology is not just its process, but also its application within a particular context.

Using a formal evaluation model would have helped both the Claims Training Evaluation Task Force Committee and Litigation Management Evaluation Task Force Committee evaluations over part of this methodological hurdle.

In particular, the Claims Training Evaluation Task Force Committee suffered for not having used an accepted methodology. For example, during the presentation of results on June 6, 1994, I was the chairman of the presentation committee. The Claims Training Evaluation Task Force Committee final report was challenged upon its

methodology with the intent of undermining the recommendations by establishing that they were of no value due to a faulty methodological process, and while I was unable to record the questions asked, my recollection and comments in an e-mail to other committee members post presentation reflect that it was a very direct challenge. In my role as chairman, it fell upon me to defend the evaluation. It was a difficult process as the methodology used was logical given the approach that the Claims Training Evaluation Task Force Committee took, but it was not sound in the face of formal evaluation models. Critics analyzing the Claims Training Evaluation Task Force Committee evaluation against the backdrop of the Provus and Stufflebeam et al. Evaluation Models should have been able to pull it apart.

The Litigation Management Evaluation Task Force Committee, on the other hand, was working with a less complex evaluation. That is, its mandate was to determine whether there was compliance with Bulletin CDB894. While it did not utilize a formal evaluation process, the evaluation process it did use was both logical and similar to what might have been used had a formal evaluation model been considered. It begs the question as to whether or not the Litigation Management Evaluation Task Force Committee intuitively used a method similar to a formal evaluation

process or whether they fell into it, as there was really no other logical method of dealing with the issue at hand. It was possible that the experience level and technical expertise of the Litigation Management Evaluation Task Force Committee members led the committee to approach the evaluation in a similar fashion to problems and events that are dealt with on a daily basis. That is, there could be a general application of existing business coping skills, coupled with common sense, and "luck" which led to the success of Litigation Management Evaluation Task Force evaluation. The "luck" component is difficult to measure or explain but the nature of the corporation's business is one that focuses upon risk taking. Risk taking relies upon working very hard at being "lucky".

Evaluation committees can learn from this examination of methodology. Similar to the importance of clearly defining the objectives of the programme being evaluated, the soundness of the methodology used is extremely important in determining the success of the whole evaluation process. The methodology section is the underpinning of the evaluation.

Determination of Results. Both the Claims Training Evaluation Task Force Committee and Litigation Management Evaluation Task Force Committee could have benefited by the realization that while it is necessary to be results

oriented, the focus of the evaluation should not be upon the results.

The results are what they are and if the evaluation process is sound the results will naturally flow from the evaluation. In other words, if the objectives are clear and the evaluation methodology is unshakeable, the results will be able to withstand criticism. If there is a weakness in an evaluation it will be in the evaluation process itself as opposed to the results.

Recommendations. Recommendations are made by examining the results and referencing the original mandate. Both the Claims Training Evaluation Task Force Committee and the Litigation Management Evaluation Task Force Committee made recommendations. The Claims Training Evaluation Task Force Committee recommendations were broad in nature and flowed from a compilation of the results of the focus groups. There was a gentle anonymity for the Claims Training Evaluation Task Force Committee members as the recommendations could be shown to be reflections of the will of a larger group.

The Litigation Management Evaluation Task Force Committee, on the other hand, made specific recommendations upon an interpretation of the results of the evaluation. This committee took full responsibility for its

recommendations. The strength of its recommendations came from directly addressing its objectives.

In this study, the recommendations of each evaluation committee were provided to the "Decision Makers" who gave the evaluation committees their original mandates. Whether or not the recommendations were "weighted" or implemented was beyond the scope of the evaluation committees and was strictly a decision made by senior management. They did, however, form a valuable piece of the puzzle for the decision makers, as would be expected, and the more confidence that could be had in the evaluation process, the more weight that could be afforded to the recommendations. The end result, though, with respect to implementation or action upon recommendations, could be influenced by factors which may or may not be known to the evaluation committee.

As is evident in both the Provus and Stufflebeam et al. Evaluation Models, the process ends for the evaluation committee with a financial analysis. This brings the evaluation process to a close from the perspective of the evaluation committees, but implies that any further evaluation will be conducted by the people responsible for commissioning the evaluation.

In the two evaluations at hand, very few of the recommendations made by the Claims Training Evaluation Task Force Committee were implemented, whereas the

recommendations of the Litigation Management Evaluation Task Force Committee were implemented in their entirety. This may have been a reflection upon the confidence that the decision makers had in the evaluations themselves, or due to other factors. In any case, however, neither of the committees was privy to the decision makers' rationale.

Developing Foundations of Evaluation

This study reveals that evaluations should consider the distribution of time and effort in conducting an evaluation. It also suggests that there should be a greater emphasis on the setting up of an evaluation rather than upon the determination of results. In other words, as Provus and Stufflebeam et al. suggest, objective clarification should be the first step in any evaluation. Consequently, objective clarification should be of primary importance as it lays the foundation for the balance of an evaluation. This is reflected in the evaluations conducted by both the Claims Training Evaluation Task Force Committee and Litigation Management Evaluation Task Force Committee.

Both the Provus Discrepancy Evaluation Model and the Stufflebeam et al. C.I.P.P. Evaluation Model emphasize the need to develop clear objectives at the outset of an evaluation. Without a clear focus at the beginning of an evaluation, the balance of the evaluation is weakened. This concept is demonstrated in both the Claims Training

Evaluation Task Force and Litigation Management Evaluation Task Force evaluations.

The Claims Training Evaluation Task Force Committee did not have clearly defined objectives and the results were questionable. The Litigation Management Evaluation Task Force Committee, on the other hand, had a specific objective and the results were solid as they could be compared directly to the objectives. The importance of having a solid foundation to an evaluation by having clearly defined objectives cannot be understated.

This study, based on two cases, seems to confirm the Provus and Stufflebeam et al. assertion that the greatest emphasis in an educational evaluation should be upon objective development and the methodological structure. By extension, this could be applied to all evaluations as demonstrated by the two corporate evaluations examined in this study.

The evaluation pyramid in Figure 1 demonstrates the proportional importance and amount of effort that should be expended in the development of an evaluation. Objective clarification forms the foundation of an evaluation. The formation of a methodological process should build upon the objectives and so on through the evaluation hierarchy. This is similar to the emphasis that the Provus and Stufflebeam et al. Evaluation Models place on objective development. If

there is a weakness at any level, then everything above that level would see that weakness magnified. As that weakness becomes magnified, it flows all the way through the balance of the evaluation and renders each higher level suspect and open to criticism. The reverse is also true; the more solid the lower levels are, the more confidence that can be had in the higher levels.

Figure 1 below illustrates my position with respect to this principle as applied to the Litigation Management Evaluation Task Force Committee.

The evaluation pyramid could be applied to both the Claims Training Evaluation Task Force and Litigation Management Evaluation Task Force Committees. Each committee passed through the various levels of the pyramid in ascending order, but with a different emphasis for each committee. The Litigation Management Evaluation Task Force Committee evaluation proceeded through the levels of the evaluation pyramid as outlined in Figure 1 below. As suggested by the Provus and Stufflebeam et al. Evaluation Models the Litigation Management Evaluation Task Force Committee placed a greater emphasis upon objective clarification, methodology and so on up through the pyramid, which resulted in a greater emphasis upon the fact based foundation and provided a solid basis for interpreting results and making recommendations.

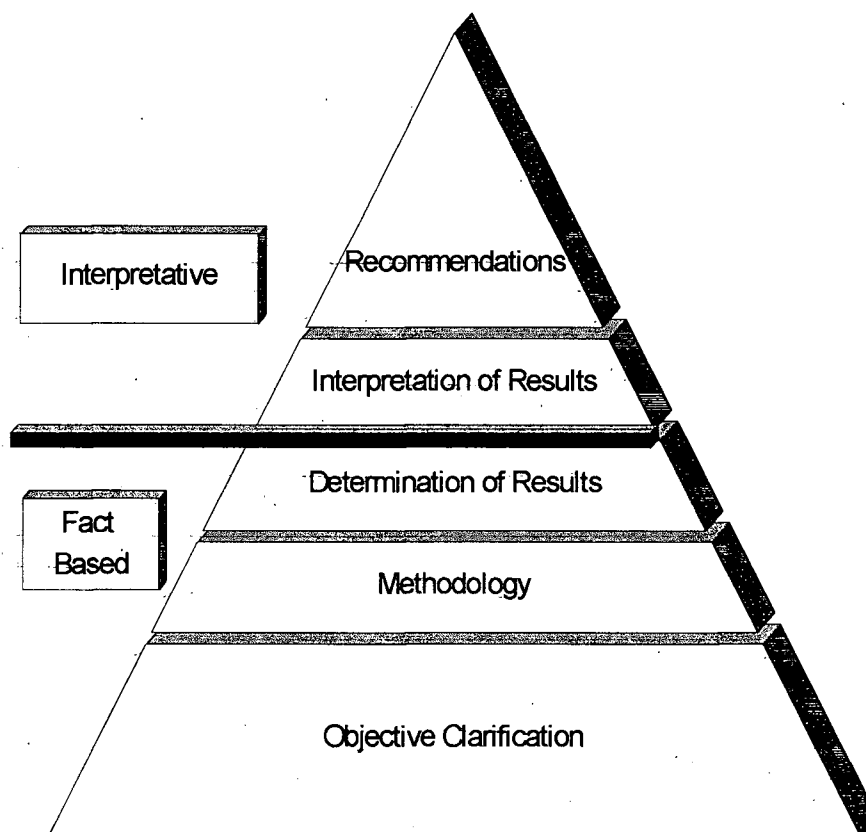


Figure 1. Litigation Management Evaluation Task Force Committee Weighting of Evaluation Processes. The amount of weight placed upon each of the evaluation stages is illustrated by the proportional blocks of the pyramid. The greatest emphasis was placed upon objective clarification which formed the foundation of this evaluation. The fact based stages are objective whereas the interpretative stages are subjective in nature.

The Claims Training Evaluation Task Force Committee placed little or no emphasis on objective clarification and chose to concentrate upon data collection, interpretation of results and making recommendations. By not concentrating upon objective clarification, contrary to the emphasis

placed on objective development by Provus and Stufflebeam et al., the Claims Training Evaluation Task Force Committee based its evaluation upon assumptions which arose from an expectation that a focus upon results would highlight problems within the Claims Training Department. By overemphasising the needs of the recipients of claims training and focusing upon interpretation of the results and recommendations, rather than identifying the objectives of claims training, the Claims Training Evaluation Task Force Committee developed an evaluation process that was methodologically unstable by theoretical educational evaluation processes. This emphasis would produce an inverted pyramid as illustrated in Figure 2 below.

Given the lesser emphasis upon objective clarification, all levels above this stage will see that weakness magnified. The result would be that recommendations made in this situation would be open to question as the foundation of the evaluation would not be solid.

The Litigation Management Evaluation Task Force Committee demonstrates this principle in a simplistic form. The Litigation Management Evaluation Task Force Committee evaluation objectives were very clear, as they were set out in Bulletin CDB894. There could be no misunderstanding as to what those objectives were, as they were a published document. The first level was therefore very solid and

provided a good foundation upon which to build the balance of the evaluation.

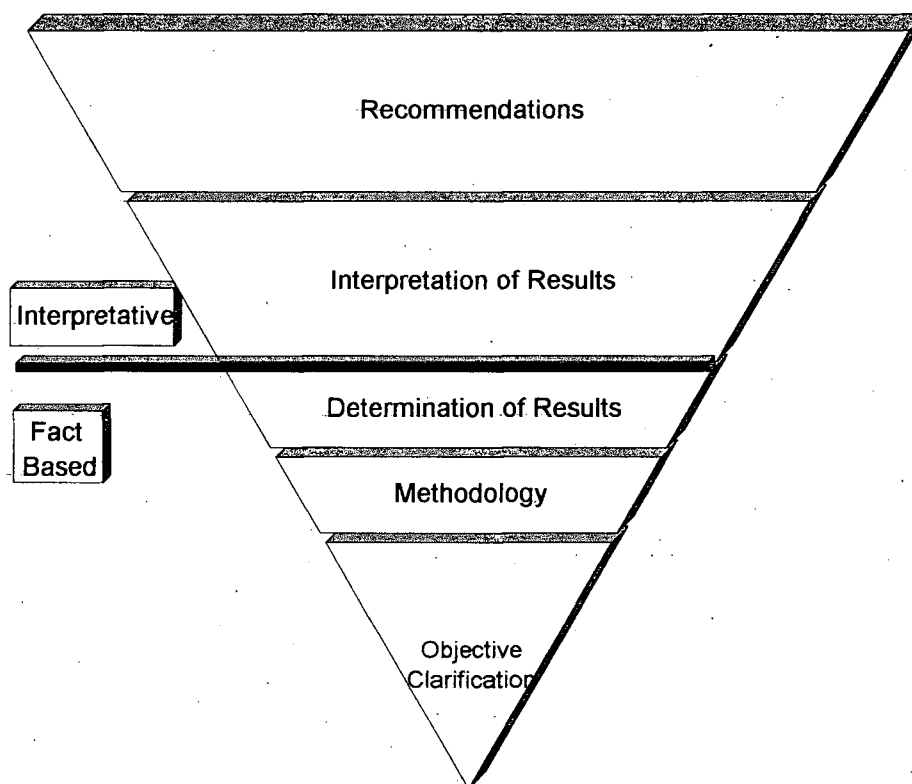


Figure 2. Claims Evaluation Task Force Committee Weighting of Evaluation Processes. The amount of weight placed upon each of the evaluation stages is illustrated by the proportional blocks of the pyramid. There was a very heavy weighting upon the Interpretation of Results but the greatest emphasis was placed upon Recommendations. Recommendations formed the foundation of this evaluation and the instability of the evaluation's methodology is represented by the inverted pyramid. The fact based stages are objective whereas the interpretative stages are subjective in nature.

The methodology used by the Litigation Management Evaluation Task Force Committee was very basic as it merely

had to establish whether or not there was compliance with the objectives. This translated into a rudimentary "yes" or "no" format. Again, the methodology consisted of asking a sample of the people involved with Bulletin CDB894 whether or not they complied with Bulletin CDB894. This is very difficult to challenge except on the grounds of sample size. That is, does the size of the sample and the method of sampling accurately reflect the larger population? The results that flowed from the application of the evaluation method are therefore quite solid.

As an evaluation proceeds from this point to a position that is higher in the pyramid, confidence in the outcomes becomes softer as they now rely upon interpretation. Challenges could arise at this stage of an evaluation with respect to how the accumulated data were interpreted as opposed to how the data were accumulated. This relationship could be the subject of future studies. This study revealed that the Litigation Management Evaluation Task Force Committee was in a stronger position than it would have been otherwise, as it had a strong base upon which to develop its interpretation. An accurate interpretation will largely depend upon the experience and skill of the committee members in interpreting results within the context of the evaluation.

Evaluation Approaches Must be Flexible in Practice

This study suggests that as a situation unfolds, more than one model may be at work at the same time in the evaluation process. That is, the Litigation Management Evaluation Task Force committee evaluation functioned as a combination of an expertise oriented evaluation along with an objectives oriented evaluation. A generalization that may be made from this study is that one evaluation model operates as the main focus of an evaluation, but other evaluation models may operate around that primary model. As the situation changes or the need arises, the focus of an evaluation alters sufficiently to obtain whatever information is necessary for the main evaluation. That is, the evaluation process may shift from one model to another and then back again. This leads to a hybrid approach to evaluation, where there is one main evaluation process and from that process, as the situation or needs change or the evaluation model in use "stalls", other models are brought in, discarded, or re-introduced. This shifting of models requires that the evaluation process being used is flexible in both direction and approach in the actual business setting. Evaluations that are formula oriented are very focussed in a specific form of evaluation, and can become static. Formulas that are fixed are not adaptable to process change.

Stall Points

The results of this study indicate that the two models under examination are effective to a certain point. They serve the need of establishing a framework for evaluation and assist in the gross data gathering process. The current study illustrates that the Provus and Stufflebeam et al. evaluation models reach a stage in the evaluation process where they become ineffective. I call this stage the stall point. This stalling point would appear to be variable depending upon the circumstances of the evaluation. In other words, given a different evaluation under different circumstances, the stall point may be reached earlier or later. The indicated conclusion, however, is that at some point these models will stall. The impact that this conclusion will have on educational evaluation thought is that evaluators, in applying models, must be cognizant of potential stall points. Rather than trying to force a situation to fit the model, one must be prepared to recognize stall points and shift from one model to another. I call this process "phase shifting". Phase shifting (or transition) occurs in order that a bridge may be developed between one model and the next. This concept was evident in the Litigation Management Evaluation Task Force Committee where the committee began in a strictly objectives oriented approach and then shifted to a management oriented approach

once the determination had been made that there was non-compliance with Bulletin CDB894. The committee continued to evaluate the litigation process after redefining its mandate but with a different focus. The significant challenge for evaluators will be to predict the potential stall points and be prepared with alternative models as evaluations proceed through model phase shifts. This study suggests that evaluators must widen their focus from a single model utilization approach to a multiple model approach. This will represent a dramatic change in the thought process with respect to the development of theoretical models. Rather than looking for justification for theoretical models and proving their efficacy through a testing process, one will now have to look at the potential of evaluation models within a spectrum of evaluation models in order to determine an holistic approach to evaluation. It will no longer be adequate to look at evaluations within the context of an approach based model. Rather, the evaluator must become a connoisseur of evaluation in order to choose the best models available within the context of the situation and apply them appropriately in order to have the best flow of evaluation throughout what will now be a continuum of evaluation.

The three main areas upon which to focus are the stall points, the phase shift or transition period and the bridging of one model to the next in order to select the

appropriate model or models which may fit potential futures. A fourth area of interest is to look at evaluation as being a continuum rather than a static process. This study suggests that specific evaluations may have start points and end points, but these are placed in a past and future context. This line of thought establishes that evaluation is an ongoing process that consists of transitions or phase shifts from one evaluation form to another with no true end point. It is more of an evolution of evaluation rather than the completion of an evaluation. By looking at an evaluation as being an evolutionary process with stall points, phase shifts/transitions and bridges, models will no longer necessarily fall out of use. Rather, their use may be changed and limited in order to serve specific purposes. New lines of thought may be developed with respect to evaluation, by revisiting and reviewing those evaluation models which were effective in the past and have fallen out of use in order to determine whether or not aspects of those models may serve a purpose or a role in current evaluation thought.

Recommendations

The interpretative stage of an educational evaluational process, as discussed in the evaluation pyramids above, is not considered by either Provus or Stufflebeam et al. The effect of the "expertise" of the evaluators in

interpretation of results by the evaluation committee would be a revealing topic for further study.

For example, the Claims Training Evaluation Task Force Committee had a broad range of cross-divisional representation and experience due to the makeup of the committee. The interpretation of the results was similarly broad, cross-divisional, and on a grand scale. This may have been due as much to the diversity of the committee members as to the scope of the mandate.

The Litigation Management Evaluation Task Force Committee was made up of members from the same division with similar technical expertise and focus. The recommendations of this committee were very specific and focussed towards addressing the objectives. The emphasis was upon evaluating Bulletin CDB894 with the aim of controlling the expense of litigation and improving the litigation process. These results may have been directly related as much to the special interests of the committee members as to the original mandate.

A recommendation for future studies would be to examine several issues.

1. Is there evidence of stall points occurring in the use of other models in other situations, or is this limited to the particular/peculiar study at hand?

2. If these stall points exist, is there evidence that other models have the ability to take over where the initial model has stalled?
3. If there are stall points and if the use of other models would appear to be appropriate in order to move the evaluation off the stall point, an examination of the transition or phase shift would be required in order to determine the dynamics of the phase shift process.
4. An examination of the models that are used to break off an evaluation stall point would be required to determine whether or not these subsequent models also have stall points.
5. If subsequent stall points also exist, it would be necessary to examine whether or not this concept of stall points and phase shifting is part of the process of an evaluation continuum.
6. Does the expertise level of the members of the evaluation committee have an effect, direct or indirect, upon an evaluation?
7. Does the exposure of career risk to the members of an evaluation committee have an effect upon the design, function and interpretation of the results of an evaluation?

These recommendations may require a long term examination of several evaluations in a variety of evaluation settings in order to establish a foundation for this theory becoming generalizable.

EPILOGUE

Impact of External Influences on Evaluation ModelEffectiveness

In discussing the personal risk of the members of an evaluation committee it is instructive to examine the impact of external influences upon the two corporate evaluations that were used in this study. By reflecting upon this background the importance of examining evaluation committee member risk is highlighted and emphasizes the need for further investigation into this unexplored facet of evaluation.

External Influences Acting Upon the Claims Training Evaluation Task Force Committee. In Spring, 1994 the then-Manager of the Human Resources Development Department (H.R.D.) was under pressure to look for ways to reduce head count (staff) within his Department in support of an overall Human Resources Division downsizing initiative. Apparently without consulting other Senior Managers or his staff he devised a plan to reduce head count within his Department. He saw the Claims Training Department, which reported to him, as an area where he could downsize. As part of the development of this plan he spent time in a field office to acquaint himself with field operations. He had little or no direct claims experience and so spent one day in a claims

office to familiarize himself quickly with the process.

This was where we first met and discussed training. He was aware that I was working on my Master's Degree in Education and wanted to discuss some ideas with me. We spent approximately two hours discussing my course work and claims training in general. He advised me of his plan and asked that I keep it confidential until it was announced. At the time, I was unaware that our conversation was anything more than a casual conceptual discussion. He later communicated this plan to the affected staff members of the Human Resources Development Department and Claims Training Department and it was at this point that a crisis developed.

The plan involved reducing the head count in the Claims Training Department which at that time fell within the Human Resources Division. This was to be achieved by decentralizing claims training, which was conducted at Head Office by Claims Training Managers, and having Field Managers assume a greater role and responsibility for training. By having less formal training conducted at Head Office by Head Office Trainers there would be less need for training staff within his Department. A core of trainers would be maintained for those areas of training which could not effectively be decentralized and the remaining trainers would be reassigned to field offices. The field offices belong to the Claims Division. By relocating staff the

Human Resources Division would reduce its head count but the resulting impact would be an increase in the head count of the Claims Division. Also, the placement of trainers, who had spent the last ten to fifteen years in Head Office, was an unsettling and very emotional issue. The Claims Division was upset for various reasons including the lack of consultation and the fact that more work would be imposed upon an already heavily loaded front-line management.

Senior management reviewed the situation and directed that a task force be created to evaluate the Claims Training Department and to make recommendations. The Claims Training Evaluation Task Force Committee would operate at arm's length from the Manager of the Human Resources Development Department. The Manager of the Human Resources Development Department selected me to be on the task force specifically as a result of our previous conversations.

The whole situation was very difficult for the members of the Claims Training Evaluation Task Force Committee as they were at personal risk from a career limiting perspective and had, as well, a vested interest in the outcome of the evaluation. The Claims Training Evaluation Task Force Committee members included a minority who were neutral with respect to the evaluation outcome and the majority were selected from front line staff who were the recipients of training, front line managers who would have

been directly affected by the original plan of the Manager of Human Resources Development Department (decentralizing training and moving it out into the field), office managers who would have been held accountable for implementing the original plan, and claims trainers who would have been directly affected by downsizing and relocation. The members had all of these concerns with them when they entered the first meeting. These concerns would be a rational explanation for why most of the first meeting skirted discussing objectives and was more focussed on the impact any changes would have on individuals.

The Claims Training Evaluation Task Force Committee could have requested clarification of the mandate and then its objectives would have been clearly defined. However, in this instance, having an open-ended and unfocussed mandate allowed the evaluation team the latitude to interpret what information would be required to satisfy the mandate reasonably and devise an evaluation process which would minimize risk. While it was never established as an operating plan, the movement directly into stakeholder interviews and focus group sessions allowed the Claims Training Evaluation Task Force Committee to deflect some of the risk by taking on a reporting role rather than an evaluation role. In other words, the Claims Training Evaluation Task Force Committee would be reporting on the

state of affairs of the Claims Training Department by recording the general views held by the Corporation through identifying the opinions of a representative cross section of the population that either had direct contact with, or received training from, the Claims Training Department. The security of the large group reduced the amount of personal risk held by the Task Force members. The Claims Training Evaluation Task Force Committee recommendations were made with a relative amount of safety knowing that they were suggested solutions to problems that were identified by a larger group rather than the results of an evaluation made by a select group of individuals.

The influence of external factors could directly influence the application of the Provus and Stufflebeam et al. Evaluation Models as well as the usefulness of the results of the evaluation. Neither Provus nor Stufflebeam et al. considers the influence of external factors in the design of their models.

External Influences Acting Upon The Litigation Management Evaluation Task Force Committee. The Litigation Management Evaluation Task Force Committee, on the other hand, did not encounter the same level of personal risk, as the impact of the evaluation would not affect people who were in a peer or superior relationship to the task force members.

A cursory financial examination of operating costs established that a considerable amount of money was being spent on defence costs. Given the public nature of the corporation, it was expected that opportunities to act as defence counsel would be made available to the private sector as opposed to creating an in-house defence bar. This situation allowed abuses in several forms to take place. Such abuses included overbilling on time spent on telephone calls and file reviews, delaying settlement or the handling of files to increase the period of billable time, transferring files to counsel within the same firm and charging for new reviews each time this occurred, and assigning minor files to senior counsel who charge at a higher billing rate and then transferring files to a junior counsel just before trial. The corporation found itself in a difficult position with respect to having to use counsel from the private sector for defence work and yet having virtually no formal control over their billing practices. "Errors" that were discovered by those claims staff who were vigilant were brought to the attention of defence counsel who usually would not resist the correction. Given the volume of defence work it was a difficult task to control these "errors". Claims Division Bulletin 894 was created, in consultation with defence counsel, to bring in a measure of control to defence counsel billing practices. It simply

required defence counsel to review the file material at the outset of the legal process and to develop a file budget for legal costs. If there was going to be an overrun, then defence counsel was required to discuss the situation with the adjuster. The adjuster was required to monitor the legal file handling from a cost perspective in addition to giving file direction.

Objections were made at the time that Bulletin CDB894, Litigation Budgeting Process, was instituted. From an adjusting point of view it was too cumbersome a process. From a defence counsel perspective it was unreasonable to expect defence counsel to be able to forecast a budget and be accountable for the accuracy of the budget of litigated files, which may take two years or more to settle or reach the courtroom.

From the outset there were warning signs that compliance by staff and defence counsel was going to be questionable. Without ascertaining that these compliance issues were dealt with prior to implementation, the intended outcomes were in jeopardy.

Bulletin CDB894, Litigation Budgeting Process, was implemented with the expectation that it would save a considerable amount of time and defence cost expenses. When these savings failed to materialize, the Litigation Management Evaluation Task Force Committee was formed to

evaluate the effectiveness of Bulletin CDB894, Litigation Budgeting Process.

From a personal risk impact perspective, the Litigation Management Evaluation Task Force Committee members were essentially risk free. They were entering the situation acting in the role of auditors. Their mandate was simply to determine whether or not defence counsel and adjusting staff were complying with Bulletin CDB894, Litigation Budgeting Process. If there was compliance, then no-one was at risk. If there was partial or no compliance, the only people on the Task Force who would have been at risk were the Centre Manager, who was leading the Task Force and responsible for the four Claims Managers in his office and staff, and the Claims Manager (the participant observer), who was responsible for ensuring that my staff complied with procedures. As the greater onus was on defence counsel to comply and non-compliance was potentially corporate wide, the amount of personal risk was minimal. Also, being on the task force would potentially lead, without necessarily direct intent, to a softening of the criticism of management and a polarization of criticism towards staff and defence counsel should non-compliance prove to be the case. I wrote the final report which was scrutinized and edited by the Centre Manager and Regional Manager, and while this was not

the case, the potential also existed for criticism in the final report to be slanted away from management.

In a similar nature but not as obvious as in the Claims Training Evaluation Task Force Committee, there was an element of bias present which could affect both the design of the evaluation and the reporting of the results. The members of the Litigation Management Evaluation Task Force Committee were selected because of their experience and commitment to litigation and the Tort (the right in law to sue another party for a wrong committed) system. This predisposed commitment could not help shaping the evaluation perspective. The personal risk, though minimal, is just another of the potential influencing factors which have no provision or consideration within either the Provus or Stufflebeam et al. Evaluation Models and is worthy of further study.

References

- Barnes Craig & Associates (May, 1993). The Request For Proposals: Abberation or Opportunity, USA
- Best, J.W. (1977). Research in Education (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Cousins, J.B. (1996). Consequences of Researcher Involvement in Participatory Evaluation. Studies in Educational Evaluation, 22, 20-23.
- Garaway, G.B. (1995). Participatory Evaluation. Studies in Educational Evaluation, 21, 85-102.
- Guba, E.G. & Lincoln, Y.S. (1989). Fourth Generation Evaluation. Newbury Park, California: Sage Publications.
- House, E.R. (1995). Professional Evaluation: Social Impact and Political Consequences. Newbury Park, California: Sage Pyublications
- Langenbach, M. (1993). Curriculum Models in Adult Education. Malabar, Florida: Krieger Publishing Company.
- McMillan, J. H., Schumacher, S. (1989). Research in Education: A Conceptual Introduction (2nd ed.). USA: Harper Collins Publishers.
- Meigs, R. F., Meigs, W. B. (1989). Financial Accounting (6th ed.). Toronto, Ont: McGraw-Hill.

- Milkovich, G. T., Glueck, W. F., Barth, R.T., & McShane, S.L. (1988). Canadian Personnel/Human Resource Management: A Diagnostic Approach. Plano, Texas: Business Publications, Inc.
- Miller, J. P. Seller, W. (1990). Curriculum: perspectives and practice. Toronto, Ontario: Copp Clark Pitman Ltd.
- Robbins, S. P., Stuart-Kotze (1986). Management Concepts and Practices (Canadian Edition). Scarborough, Ontario: Prentice-Hall Canada, Inc.
- Stanley, J. C. & Hopkins, K. D. (1972). Educational and Psychological Measurement and Evaluation. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Tyler, R. (1986). Changing Concepts of Educational Evaluation. International Journal of Educational Research, 10, 53-55.
- United States Automobile Association (1988). Litigation Management. USA
- Wolf, R.M. (1987). Education Evaluation: The State of the Field. International Journal of Educational Research, 11, 3-6.
- Workplace Training Systems Open Learning Agency. (September 1993). Corporate Training Consortium Initiative: Managing Diversity Videoconference Cost-Benefit

Analysis of Decentralized Vs Centralized Delivery

Approaches. Vancouver, BC

Worthen, B. R., Sanders, J. R. (1987). Educational

Evaluation: Alternative Approaches and Practical

Guidelines. White Plains, N. Y: Longman.