A STUDY OF THE APPLICATION OF A MODIFIED TRANSACTIONIST PROGRAM EVALUATION MODEL TO A NON-FORMAL DISTANCE EDUCATION COURSE FOR COMMUNITY-BASED ADULT LITERACY PROVIDERS

CENTRE FOR NEWFOUNDLAND STUDIES

TOTAL OF 10 PAGES ONLY
MAY BE XEROXED

(Without Author's Permission)

BLAIR W. KETTLE

National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

A STUDY OF THE APPLICATION OF A MODIFIED TRANSACTIONIST PROGRAM EVALUATION MODEL TO A NON-FORMAL DISTANCE EDUCATION COURSE FOR COMMUNITY-BASED ADULT LITERACY PROVIDERS

by
Blair W. Kettle, B.A., B.Ed., Grad. Dip. Ed.

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Education

Division of Learning Resources
Faculty of Education
Memorial University of Newfoundland
May 1994

St. John's Newfoundland

Abstract

The purpose of this study was to select and apply an appropriate evaluation methodology for the Distance Education for Literacy Providers Course, a federally and provincially funded pilot project designed to help community-based adult literacy tutors acquire skills and knowledge that would help them in their work. The review of related literature provides background information about distance education and educational program evaluation, and serves to support the selection of the transactionist evaluation model which was used for the study.

Robert E. Stake's Responsive Evaluation approach was used as the guiding methodology for the study. It was chosen because it is the most widely known and tested transactionist methodology, and because it offered flexible, rigorous, and context-sensitive methods for such things as audience identification, concerns and issues identification, and standards development.

Qualitative data were gathered through on-site observations, interviews, document analysis, and questionaires containing a blend of forced-choice and open-ended questions. Data collected was compared to the evaluation standards so that evaluation judgements could be made.

The study concludes with recommendations made about the Course, and the limitations of the Responsive Evaluation approach.

Acknowledgements

I wish to express my appreciation to those people who were associated with, supported, and encouraged me in completing this study.

My greatest debt is to Dr. Mary F. Kennedy. Her endless energy, dedication, and patience kept me on track when I lacked these things myself. I feel honored to have been her student.

I wish to acknowledge the assistance of Genevieve Gallant for her work in gathering data for the study. Thanks also go to the developers and the participants of the Distance Education for Literacy Providers Course for their co-operation.

Special thanks go to my wife, Jill Marshall, for her love and patience, and for putting up with the constant mess of paper and books around the computer.

Finally, thanks go to my parents, Nathan and Patricia Kettle, and to my parents-in-law, George and Joan Marshall, for their love, encouragement and financial assistance in my lean times.

TABLE OF CONTENTS

Abstract ii

Acknowledgements iii

List of Figures viii

CHAPTER 1 1

Background to the Study 1

Introduction 1

The Distance Education for Literacy Providers Project . . 1

Significance of the Study 4

Limitations of the Study 5

Organization of the Study 6

CHAPTER 2 8

Review of Related Literature 8

Introduction 8

Distance Education 8

Origins 8

The industrial revolution and education. . . . 14

Emancipation through education. 16

The private correspondence school 20

From Correspondence Study to Distance Education . . 22

A new image for distance education. 26

Keegan's typology. 28

Autonomous Distance Teaching Institutions 30
Dual Mode Institutions 31
Evaluation Within Distance Education 38
The Economics of Conventional Education 39
The Economics of Distance Education 40
Learner-centred Vs. Institution-centred Education . 42
Equalizing Autonomous, Dual Mode and Conventional Systems 43
Program Evaluation 44
Program Evaluation in Historical Perspective . . . 46
Nineteenth Century Quest For Improvement 47
Horace Mann. 47
Egerton Ryerson. 48
The Evolution of Evaluation Methods 49
A focus on objectives. 51
Industrialization, World War II, the post-war era. 51
A Taxonomy of Evaluation Models 55
The Major Evaluation Models 58
Systems Analysis 58
Advantages. 62
Disadvantages. 63
Behavioral Objectives 65
Advantages. 67
Disadvantages. 68

Decision Making 70
Advantages. 71
Disadvantages. 72
Goal Free 73
Advantages. 74
Disadvantages. 75
Art Criticism 77
Advantages. 78
Disadvantages. 79
Accreditation 80
Advantages. 82
Disadvantages. 83
Adversary 84
Advantages. 86
Disadvantages. 87
Transaction 88
Advantages. 93
Disadvantages. 96

Chapter 3 98
Methodology 98
Evaluation Model Selection 98
Methodological Procedures 102
Audience Identification and Consultation 103
Concerns and Issues Analysis 104

Document Analysis 105
Evaluation Standards 105
Observation Data Collection 111
Evaluation Instruments 112
Data Analysis 114

CHAPTER 4 116
The Results of the Study 116
Introduction 116
Interim Report 116
Final Report 132
Descriptive Data 136

CHAPTER 5 146
Summary, Conclusions and Recommendations 146
Summary 146
Recommendations 146
Conclusion 152
References 155

APPENDIX A - Indepth Interview 173
APPENDIX B - Evaluation Instruments 188

List of Figures and Tables

Figure

1 A taxonomy of major evaluation models 57

2 Prominent events in a responsive evaluation 93

3 An adaption of Stake's prominent events in a responsive evaluation 102

Table

1 Level of satisfaction with curriculum 117

2 Level of satisfaction with program by training 117

3 Frequency of attendance 118

4 Program objectives 121

5 Participant interaction at teleconference sessions 123

6 Participant rating of self-confidence 125

7 Perceived benefit derived from program 127

8 Suitability of teleconference system 129

9 Course aspects cited as being most beneficial 131

10 Course participants' tutoring history 133

11 Level of tutoring activity 134

12 Effects of course on tutoring activity 134

13 Participants' rating of course 136

CHAPTER 1

Background to the Study

Introduction

The increasing emphasis being placed on community and regionally appropriate educational development programs, particularly those delivered through a distance technology mode, has necessitated finding an effective and practical means of determining the merit and worth of such programs. It is with this purpose that this study was initiated.

The study examined the various models for program evaluation as set out by House (1978) in his Taxonomy of Major Evaluation Models with a view toward identifying a methodology appropriate for application to any small to medium scale community-based distance education programs for semi-professional development. The study took place in Newfoundland, and focused on a federally and provincially funded pilot project called the Distance Education Course for Literacy Providers.

The Distance Education for Literacy Providers Project

The Distance Education for Literacy Providers (DELP) Course was initiated as a pilot project, utilizing the technical resources of Memorial University's Telemedicine and Educational Technology Resources Agency (TETRA) to deliver a distance mode educational development program to adult literacy practitioners in the volunteer, college, and community-based sectors of

Newfoundland and Labrador. The DELP Course began from an idea discussed between a professor and specialist in reading from Memorial University's Faculty of Education, a specialist in adult education from the Division of Extension and Continuing Studies, and an adult literacy practitioner and social advocate with many years of experience in community literacy development. It was developed with grants from the Government of Newfoundland and Labrador and the Secretary of State of Canada.

The project involved developing a non-credit course which could be delivered with the aid of a province-wide teleconferencing network that could connect to any site on the Island or in Labrador where telephone services are available. The intent was to develop a course for literacy practitioners which was accessible, regardless of geographic location.

The principal instructional media for the DELP Course were teleconferencing, video and print. The print and video materials were mailed to participants in advance. At each of ten consecutive weekly two-hour teleconference sessions the teleconference facilitator would lead participants through pre-assigned activities and generate discussion on each week's agenda topics. Most of the weekly sessions also included invited guest speakers who would talk and answer questions on topics related to adult illiteracy and literacy development.

The DELP Course was intended to serve a wide audience of literacy providers. Thus, the pilot offering included both volunteers (some of whom who had previous training in adult

literacy development), and professionals in the field of adult education. The latter group worked primarily in the community college system.

A condition of funding for the DELP Course required that it be evaluated by an external evaluator and that a copy of the evaluator's report be sent to the funding agencies. In the Call For Evaluation Proposals, the DELP Course developers specified that evaluators collect data on the following:

1. the ability of participants to utilize program content in their literacy practice.
2. the effectiveness of each medium, and the program design.
3. the duration of the program.
4. the active engagement of participants in the learning process.
5. the future utilization of program materials as useful resources for adult literacy groups.
6. the learning which could be attributed to the program. A post-program assessment of change in knowledge, attitude and skills were to be determined 6 months after program completion.

This study represents the work of the evaluation team, and includes all the recommendations which were included in the evaluation report.

Significance of the Study

The significance of this study derived from three main areas. First, it was intended that the evaluation work on which the study was based should be an effort to replicate as much as possible the evaluation methodology used by Lertpradist (1990) in *A Study of the Application of a Selected Evaluation Methodology in an Extension Setting*. Lertpradist applied a modified version of Robert Stake's "responsive approach" to the evaluation of a Thailand Government Fisheries Extension Program in fish farming developed by that Government's Department of Fisheries. Her objective was to select a rigorous evaluation model that was practical and yet consistent with the spirit and setting in which the Extension Program was offered. The success of that evaluation provided impetus for testing the validity, reliability and viability of the evaluation model for use in other settings.

This study was also significant for its attempt to find a rigorous evaluation model that would be consistent with, and supportive of, the spirit and setting in which the DELP Course was to be offered. One of the desired outcomes of the DELP Course was that community-based literacy providers should take ownership and refine and promote it to maximize its usefulness for their purposes. Nevo (1986) describes this as the socio-political function of evaluation. The evaluators were therefore concerned the evaluation model be appropriate for the sense of social empowerment engendered by the DELP Course, that it be flexible

enough to use for a program that was offered through distance education technologies, that it be congruent with the socio-political function of the Course, and that it demonstrate a high level of investigative rigor.

The study, and the evaluation on which it was based, is equally significant for its attempt to determine whether the DELP Course was effective in relation to the standards and criteria that were agreed upon by the various audiences. The ability of the evaluation approach to satisfy and balance the demands of program funders with the socio-cultural and socio-political context in which the program was developed, and at the same time provide a report that was rigorously derived and credible, was the cornerstone which marked the foundation of the study.

Limitations of the Study

The DELP Course was designed as a pilot project. It was intended that general feedback from program participants, observations on the part of course facilitators, and results from the evaluation report would be used to refine the DELP Course for subsequent offerings. For this reason, the evaluation findings must be tempered with the knowledge that an evaluation performed on later offerings of the DELP Course could produce notably different and significant findings. A more accurate reflection of the DELP Course's ability to satisfy the needs of the participants would likely result after it had been delivered

several times. However, the program funders had required that an evaluation be completed on the first implementation.

A second limitation of this study was that it used only one model to evaluate the program. It is impractical to employ multiple models for a single offering of any program, and since the program was being offered only as a pilot project, the opportunity did not exist to apply a different model to future offerings.

Organization of the Study

The evaluation of the DELP Course was organized around two evaluation reports. The first was an interim report which included a description of the DELP Course, the audiences, the evaluation process and methodology, audience concerns and issues, the evaluation standards, with respective criteria, and a conclusion which presented findings on standards which could be measured during, and within one month of the last day, of the DELP Course. The second and final report, presented six months after the first, gave a brief description of the evaluation plan, and included the findings from the standards that the evaluators were unable to measure at the conclusion of the DELP Course.

Each report concluded with recommendations that were derived from the evaluation findings and the evaluator's overall subjective assessment of the DELP Course.

This thesis is organized into five chapters. Chapter 2

presents a review of the literature on distance education, quality control in distance education, and educational program evalation. Chapter 3 relates the literature on distance education and evaluation from chapter 2 to the selection of an evaluation model. Also included in chapter 3 is a description of the methodology that was used to implement the study. Chapter 4 presents the evaluation results, and chapter 5 describes evaluation conclusions, and recommendations for program improvement.

CHAPTER 2

Review of Related Literature

Introduction

The basis for this study was an evaluation of a non-credit course which had been designed for and was offered through a distance technology mode. The purpose of the study was to outline a flexible but rigorous approach for evaluating community and volunteer-oriented courses and workshops which are increasingly being offered through the distance education mode. With this purpose in mind this chapter addresses three background topics. It first discusses the concept of distance education from a broad historical perspective and attempts to show that its origins are mirrored in the educational activity this project set out to evaluate; that is, non-formal instruction which reaches into the community and attempts to empower community leaders and volunteers for human capital development. The remainder of the chapter is dedicated to providing an overview of evaluation within distance education, and finally a more specific discussion of evaluation and evaluation models.

Distance Education

Origins

It can reasonably be argued that the contemporary practice

of distance education is simply a natural extension of several thousands of years experimentation with various forms of human communication. That is to say, the idea of educating people from a distance would have existed well before terms like correspondence study, external study, independent study, home study or distance education were ever used. Hence, when authors like Holmberg (1986) or Verduin and Clark (1991) say that correspondence study and distance education can be traced at least to the latter part of the 19th century they are using a narrow definition of the concept.

If we put aside, for the moment, two motifs that have dominated popular and traditional ideas about education, namely, teacher and classroom, the distinction between education and distance education narrows. The elements of the educational process that remain are the learners, the educational message, and the educational media. These three elements form the essence of distance education, since it is tied to neither the concept of a teacher or a classroom. Viewed in this light, it should be fair to say that distance education dates back at least 35,000 years to the time when pre-historic peoples began etching and painting the walls of caves and cliff facings. While theirs was indeed a primitive method of communicating - possibly to the gods of the pantheon or to other migratory peoples - it was effective in that it transcended barriers of geography and time.

If we consider 35,000 year old cave paintings as artifacts of communication indicative of pre-civilization, then myth and

folklore, which were the first universally effective educational and socialization media (Turner, 1979), could fairly be seen as a marker of an emerging civilization. Indeed, since pre-historic cave paintings demonstrate that learners and social, religious and educational messages predated an effective means of communicating those messages on a large scale, it is reasonable to conclude that the tools of mass communication in use at a given time are a reflection simply of the evolutional stage of the civilization of the time. Hence, it should also be reasonable to view myth and folklore - media which were able to communicate culture, "the symbolic, intellectual construction of a people" (Turner and Smith, 1979, p.386), within nomadic societies - as the foundation for an educational model that is part of the genealogy of modern distance education.

The argument here is that distance education has genetic links to the oral tradition. It is not simply, as some might suggest, a relatively modern and impoverished scheme for education which has been concocted by well-intentioned but misguided academics. Rather, it is a distant offspring of an educational and social communication process that had evolved to meet the needs of early civilization. Viewed in historical context, it should be ascribed as much legitimacy as other phenomena of civilization, such as government or money.

While oral tradition stands as a significant marker in the evolution of human communication and while it continues to operate within families and cultures (Caplan, Choy and Whitmore,

1992; Hess and Azuma, 1991), an even more significant historical marker emerged with the advent of written code (semiotics). Innis (1951) contends that the major determinants shaping the history of Western civilization have been physical communication media. He says that the span of time between the use of the stylus and the cuneiform script in ancient Mesopotamia and the development and use of broadcast communications in modern civilization has been marked by cycles of monopoly of knowledge.

> The relation of monopolies of knowledge to organized force is evident in the political and military histories of civilization. An interest in learning assumes a stable society in which organized force is sufficiently powerful to provide sustained protection. Concentration on learning implies a written tradition and introduces monopolistic elements in culture which are followed by rigidities and involve lack of contact with the oral tradition and the vernacular (Innis, 1951, p.4).

The movement from a society's reliance on an oral system of communication to one based on the less transitory nature of written language represents a significant shift in social epistemology. The ability to read necessarily includes a biased understanding of how knowledge is acquired and verified. It engenders a unique appreciation of the power of the written word, and more particularly, the ability to communicate through a

written code.

The Greeks were the first to cultivate advanced theories and systems of education, all of which were based on and transmitted through the written word. Plato, for example, illustrated through his Doctrine of Ideas his belief that man's emancipation could only be achieved through his power of understanding and that in the ideal state the level of an individuals understanding (and thereby his level of freedom) would be established through the state educational system (Reese, 1980). If we juxtapose the Platonic position with Innis' contention that a concentration on learning introduces monopolistic elements in culture which involve a lack of contact with oral tradition, it would seem an obvious conclusion that human subjugation will be the likely fate for those whose only means of communication, in a literate society, is based in oral tradition.

Oral transmission of ideas and knowledge, when compared to a written form, is an unreliable means of communicating. The more an idea is exchanged verbally, the more it is likely to become confounded (Lewis and Nichols, 1965). Unless it is intentionally changed, however, an idea that is communicated through text will remain true to its originator. Unfortunately, when it is limited to that which can be produced by hand, communication through text is necessarily an inefficient means of disseminating ideas or knowledge to a large audience because, in the absence of technology like photocopiers, it is time consuming to produce. Thus, the value that hand produced text would have for a culture

based on oral communication would be marginal at best. Only a small minority within the culture would benefit. Moreover, because accessibility to written text, prior to the invention of the printing press, was extremely limited, education became "the possession of a special class" (Innis, 1951, p.4). Not coincidentally, political power was also aligned with this class.

While the invention of the printing press eventually contributed to a new world order it did not, by itself, effect immediate and dramatic changes in the rate of literacy. In fact, literacy remained a possession of a minority for several centuries after the printing press had been developed. The catalyst for change came, ultimately, in the form of populist writers. It is a fundamental principle that technology (in this case, communication technology) is useless until an informed, creative and technically skilled mind can make it function.

Understanding that human ideas can be communicated through text and that the integrity of those ideas can be maintained, unlike ideas that are communicated orally, is essential to an understanding of the human-communications-through-technology equation. To understand this is to understand the inherent purpose and potential of text as a communication medium. The invention of the printing press made this knowledge even more poignant because, by broadening the potential audience, it gave added value to the equation. Increases in the availability of books stimulated an increase in the rate of literacy which in turn stimulated the demand for even more printed material. Thus

the skill of communicating through text increasingly became both a prized and feared possession.

The industrial revolution and education.

The Industrial Revolution - a social, economic and political upheaval which began during the mid-eighteenth century (Toynbee, 1967) represents a major turning point in the history of modern civilization. It was an era of development, of problem-solving, and of refinements in technology, government, philosophy and law. It was characterized by an accelerated growth in the principles of competition, mechanization, mass production and mass distribution. And it brought western civilization into the modern age.

Historians say that while the Industrial Revolution was nested in England, it developed rapidly and by the beginning of the nineteenth century had moved into most of the northern and west European countries and the United States (Cipolla, 1978). Moreover, recalling that the purpose of this discussion is to trace the development of distance education, it is important to focus on the fact that until the mid-eighteenth century, Europe was still an agrarian society where peasant classes worked under a feudalistic socio-political system and where government was by monarchy. The Industrial Revolution became the historical marker for a metamorphic change in the political hegemony of the time. It marked a decline in the number of agricultural workers, and a

subsequent increase in industrial labour (Cipolla, 1978; Toynbee, 1967). The result was a new social dynamic which was highly antagonistic toward the governing aristocracy and which favoured an egalitarian world order characterized by republican government.

> The new class of great capitalist employers made enormous fortunes, they took little or no part personally in the work of their factories, their hundreds of workmen were individually unknown to them; and as a consequence, the old relations between masters and men disappeared, and a "cash nexus" was substituted for the human tie. The workmen on their side resorted to combination, and Trades-Unions began a fight which looked as if it were between mortal enemies rather than joint producers (Toynbee, 1967).

The Industrial Revolution had, in effect, created social fissures where previously there were none and it widened those that already existed. All classes were assessing the foundations of the social structure. Those at the top of the social ladder had become the hereditary captains of ships overloaded with centuries of accumulated wealth; those at the bottom represented an approaching tidal wave of countless generations of peasants who had been subjugated by ignorance and militaristic force.

Emancipation through education.

During the 1770s and 1780s England saw the emergence of self-education societies. These small, grassroots organizations were energized by enlightened individuals who saw power in organization. Prior to the nineteenth century, formal education was available to an elite few. The self-education societies took as their mandate the release of the popular masses from the ignorance caused by illiteracy.

The philosophy of the ruling aristocracy was that by keeping the lower classes in ignorance they were much easier to rule. One late eighteenth English parliamentarian, for example, in response to a proposal to set up Sunday schools for working class children said, "I allow no writing for the poor...My object is not to make them fanatics, but to train the lower classes in habits of industry and piety" (Simon, 1960, p.133). Even more revealing of the paternalism characteristic of the time was a comment made by the president of the Royal Society during the 1790s; he argued that giving education to the poor was to prejudice

> [t]heir morals and happiness; it would teach them to despise their lot in life, instead of making them good servants in agriculture, and other laborious employments to which their rank in society has destined them; instead of teaching them subordination, it would render them factious and refractory, as was evident in the manufacturing

> counties; it would enable them to read seditious pamphlets, vicious books, and publications against Christianity; it would render them insolent to their superiors; and in a few years, the result would be that the legislature would find it necessary to direct the strong arm of the law against them (Simon, 1960, p.132).

In social terms the Industrial Revolution became a period of enlightenment. The peasant and working classes and the emerging middle classes had endured the hardships of feudalism and the early period of industrialization and emerged with the spirit to fight (sometimes violently) for social equity. Their symbols of emancipation became the American War of Independence and the French Revolution and their inspiration came from populist writers like Thomas Paine. Paine was a shrewd and prolific wordsmith who sold hundreds of thousands of pamphlets to a population hungry for social justice. "[B]reaking with time-honoured traditions of political writing [Paine forged] a new language that would reach out to a mass audience"; he railed against government by monarchy and for "equality of rights among all citizens" (Foner, 1988, p.9).

Donald (1983) argues that the history of literacy in England is tied directly to the pre-nineteenth century ruling class and its relationship with the "national popular masses". He argues that

> [t]he publication of Paine's Rights of Man in March 1791 stands emblematically at the start of that history. "From above", the perspective of a bourgeoisie politically nervous in the wake of the American and French revolutions, it seemed to embody what Webb, in his classic study The British Working-Class Reader 1790-1848 (1955), called "the challenge which a literate working class presented to its betters" (p.VII). "From below", in contrast, Paine's combination of popular liberalism, which was atheistic, republican, democratic and fiercely anti-aristocratic, with an assertion of the rights of the "free-born Englishman" appeared as a powerful point of ideological and political cohesion (Donald, 1983, p.36).

Donald (1983) goes on to say that in 1793 Paine's book was banned as seditious libel and that Paine himself was forced into exile. In 1795 the state put limits on the holding of public meetings as a means of causing the demise of the "radical self-education societies which had flourished during the previous decade" (Donald, 1983, p.37). One of most repressive moves, however, came in 1799 with the Corresponding Societies Act which outlawed self-education societies altogether and which put strict controls on the printing trade (Donald, 1983).

These measures, however, were simply the last gasps of a drowning aristocracy because, by the first quarter of the nineteenth century, England had charted a course for a new

'openness' toward public education. It was in this new spirit of educational reform that correspondence education was born. Verduin and Clark (1991) point out, for example, that the person generally acknowledged as being the first distance educator (Isaac Pitman) began teaching shorthand by correspondence in Bath, England in 1840. Indeed, at about this time courses offered through correspondence began emerging all over Europe and by the turn of the century colleges and universities throughout both Europe and the United States were offering correspondence courses leading to external degrees (Holmberg, 1986)

Prior to the mid 1970s distance education was known and practised, nearly exclusively, as correspondence study. It was tuition by mail and it had changed little in the century and a half since it had begun. Courses were typically developed as a series of self-study units contained in text-based material supplied by the institution. Student evaluation typically consisted of one or more assignments (sometimes supplied in advance and sometimes sent to the student at designated dates) which were completed and sent to the school for marking (Sweet, 1991). And, while tremendous advances had been made in communication technologies since the early days of correspondence study, the original tuition by mail model is still being used today (Hope, 1986).

The private correspondence school.

The image of unskilled and unemployed men and women earning trades through correspondence study and subsequently developing lucrative careers has taken its place as a classic among other motifs of the American (and Canadian) success story. And the company which probably did more than any other to create this image, contrived though it may have been, was International Correspondence Schools (ICS).

Begun in the 1890s in Scranton, Pennsylvania, and still an active element in correspondence education, ICS had its roots in the teaching of vocational and technical skills (Verduin and Clark, 1991). It developed as a response to the need for instruction that was not tied to the costly constraints associated with more traditional, locality-centred, educational and training institutions. Admittedly, many non-commercial, university-based correspondence programs were also being offered from about 1880 onward. However, institutions like ICS became, and for nearly a century remained, the torchbearers for correspondence study because their reason for existence was financial profit. They were (and are) in the business of product development and marketing.

Peters (1983) has noted that

> [f]rom the start, distance study has a special relationship with the industrial production process insofar

> as the production of study materials in itself is an industrial process built into the whole teaching process as a constituent part, quite unlike the production of text books, for example. In the case of commercial distance teaching establishments the further question of selling the printed or otherwise duplicated study units adds calculations of applied economics to the teaching process (p.97).

Riding the wave of the Industrial Revolution, institutions like ICS emerged to provide industrial style solutions to an educational problem, and to fill a commercial niche. In effect, they made correspondence study a mail-order industry and with effective advertising they harvested profit from the great American dream of career and independence. Unfortunately, for commercial correspondence schools the ratio of academic robustness to concern for profit was the inverse of what it would have been at non-profit colleges and universities. Hope (1986) points out that the free enterprise system typically provides the best and worst of any product or service. Private sector correspondence education, however, has never been a highly competitive industry and competition is typically the key to quality. Thus education by correspondence became a widely known but not widely respected mode of study.

From Correspondence Study to Distance Education

Throughout the 1960s the whole field of education in North America underwent a revival. But the germination for this revival actually began during the 1950s. In 1954, the great American psychologist B.F. Skinner published "The Science of Learning and the Art of Teaching", a seminal article which advocated the use of behavioural reinforcement theory in education. Skinner proposed the use of a highly structured and systematic model of teaching that portrayed learning as the product of linearly sequenced steps, each of which is followed by "immediate and frequent reinforcement of the learner" (Seels, 1989). His theory that the consequence of an act is the most important determinant in controlling the act itself led to the development of programmed instruction (Thomas, 1985).

Programmed instruction became the first system of instruction based on a theory of learning (Seels, 1989) and a triumph for the field of psychology (Hornstein, 1992). Since the early part of the twentieth century, the discipline of psychology had been struggling to establish itself as an empirical science (Osborne, 1991; Smith, 1992), attempting to distance itself from its more philosophical (Thomas, 1985; Hornstein, 1992) and theological (Reese, 1980) origins. The preliminary success of the high profile experiments in computer-based programmed instruction, such as the Programmed Logic for Automated Teaching Operation (PLATO) project which began in 1960 and the Stanford

CAI project which began at about the same time was a form of validation for the behaviouristic theory of learning (Saettler, 1990). PLATO became just one example which provided empirical evidence that the natural science methodology could be applied as a technology to study, quantify, test and control learning behaviours. The most comprehensive test of the programmed instruction concept, however, was in military training. Indeed, many of the now classic names associated with applied instructional systems got their start in developing training for the United States Air Force and Military (Gagne, 1989).

In historical terms B.F. Skinner's notion of a 'Science of Learning' and its offspring - programmed instruction - were derivatives of positivism which, as Gough (1988) points out, was itself an outgrowth of a generalized belief in progress that characterized the Industrial Revolution era. Indeed, as Hamilton (cited in Goodson, 1988) has observed,

> [b]y the 20th Century, the batch production rhetoric of the 'classroom system' (for example, lessons, subjects, timetables, grading, standardization, streaming) had become so pervasive that it successfully achieved a normative status (p.6).

The practice of distance education, prior to the late 1960s, was emancipatory and progressive but like other educational endeavours of the time, it was not based on a theory of learning.

After the rise of the programmed instruction movement in the 1960s, however, a new generation of behaviourist educators believed their new science could be applied (with dramatic results) to the industrial style educational problems represented by independent study.

The growth in scientifically derived principles of learning during the 1950s and 1960s was opportune. The United States was spending heavily in research and development for education in its 'race for space', and learning technologies which held the promise of a higher yield for the educational dollar attracted attention. Another important factor which kept education in the limelight was the post-war baby boom. The stress on the educational system accentuated the need for cost-effective educational expansion and radically new approaches for achieving that goal. Distance education was poised to set the pace for the educational renaissance.

Surprisingly, however, the growing courtship between correspondence study and the new behaviourally-based learning technologies did not consummate first in the United States. Rather, as Holmberg (1986) points out, the new era for distance education began with the founding of the British Open University in 1971. He says that prior to that time, virtually all large scale distance teaching organizations were private correspondence schools.

The British Open University established itself as the first large-scale educational institute dedicated to the use of the new

learning theories for the development of sophisticated, high-quality correspondence study courses. It employed a broad range of media and placed emphasis on interaction between the student and course tutors. Clennell, Peters & Sewart (1983) give the following succinct description of the Open University approach:

> The correspondence units are closely related to radio and television broadcasts and the whole teaching package of written and broadcast material is produced by a course team which is composed of a central and regional academic staff, members of the BBC and educational technologists. The student must respond to this teaching material in an active way, by carrying out experiments, writing essays, working through problems, projects, etc., and while some of this work may be used for self-assessment, the majority is assessed in written form either by a tutor (tutor-marked assignments) or by the computer (computer-marked assignments) (p.327).

The Open University became the model for other large scale distance teaching institutions around the world, many of which were established shortly after. It became the testing ground for a broad range of educational media which were capable of "harnessing industrialized processes to education" (Keegan, 1991, p.3). But the catalyst that made these media especially effective was instructional design - a systematic method for the developing

instruction which traces its origins to B.F. Skinner. The Open University, through a team approach to course production (Harry, 1990), incorporated instructional design and the growing diversity of educational media and thereby popularized the concept and the emerging field of educational technology. It revived the spirit of egalitarianism in education (Perry, 1990) which had languished in a creative void since the early part of the Industrial Revolution era and it gave new vitality to the notion of applying scientific principles of human learning, i.e., instructional technology.

A new image for distance education.

Prior to the advent of the British Open University, correspondence study was draped in connotation. It implied a specific type of learning and instructional method; it implied certain type of course subjects; it implied a certain type of learner; and it implied a certain level of educational quality. Correspondence study was paradigm-bound and it became a major hurdle to convey it through a new paradigm; namely, that represented by the term distance education. As Harry (1990) says, "[t]he first undergraduate courses [at the Open University] were presented in 1971 in the face of a considerable amount of scepticism and even hostility from the educational world" (p.15).

"For long the cinderella of the education spectrum, distance education emerged in the 1970s with a changed image" (Keegan,

1986, p.3). Unfortunately, the new image is better described as a potpourri than a clear portrait. "No single approach dominates. Current practices range from the 'deluxe' model which is the creation and hallmark of the Open University of Great Britain to a minimally structured process" (Seaborne and Zuckernick, 1986, p.37).

In early forms of distance study (i.e., correspondence education) the student worked at home, usually studying printed material provided by the institution. The use of educational media, other than text, was limited and sporadic. Educational radio was in widespread use, particularly in the United States, during the 1920s and 1930s, but its popularity waned in succeeding decades (Leach, 1983). The new era in distance education, however, brought with it new and more systematic approaches to designing instruction and new ideas about student support systems. The new thinking was that the optimal use of state-of-the-art audio-visual and communications media would facilitate distance learning and bridge the communication barrier that traditionally existed between the teacher and the learner. New thinking on course and curriculum design suggested that well designed courses would compensate for the lack of such things as "personal interest, friendly direction, timely approval, encouraging comment, leading question, concerned intervention and subtle reward" (Batey and Cowell, 1986), which are more characteristic of the conventional classroom. Smith and Small (1982), two of a new and growing group of theorists note, for

example, that while distance education will always require a significant amount of independent study, it need not be a totally reclusive pursuit. This is not to suggest, however, that all institutions which offered courses through a distance mode were enticed into the new instructional technology movement.

Mcinnis-Rankin and Brindley (1986) point out that many institutions have adopted the goal of making the distance learner as independent as possible. For these institutions this means that when a course 'package' is delivered to the learner the largest part of their responsibility ends. Other institutions, however, go to great lengths to ensure that the learner has support through a communication link with the institution or its affiliate. The point to be made here is that what was new in distance education was not only a body of theory but a growing eclectism of practice as well. Just as in conventional education, different schools of thought have led to different schools of practice.

Keegan's typology.

As the diversity of practice in distance education increased the consensus of what was actually meant by distance education declined. Observing this problem Keegan (1986), in his classic text The Foundations of Distance Education, developed a typology of distance teaching systems as a means of classifying the field. Notably, one of the premises that he used to develop his typology

Notably, one of the premises that he used to develop his typology was that

> it should only include those distance teaching institutions or departments of existing institutions which exhibit both the major characteristic subsystems of distance institutions (course development and student support services) - for without this limitation the variants are legion. Institutions or departments which are considered not to exhibit both these operational subsystems are excluded (though some of them have made excellent contributions to distance education) (Keegan, 1986, p.136).

Keegan (1986) points out that fundamental to his classification of distance teaching institutions is a basic distinction between autonomous distance teaching institutions and those which were divisions or departments within conventional teaching institutions, namely, mixed or dual mode institutions. The distinction is significant because it divides the two major schools of thought about distance education; that is, instructional systems which serve a large audience using economies of scale versus a system which offers instruction to a small to medium scale audience frugally, using campus-based resources.

Autonomous Distance Teaching Institutions

Keegan describes two types of institutions under the heading of 'autonomous'. The first type, he says, are the public and private correspondence schools and colleges. The second type is distance teaching universities or open universities. Both of these types of institutions are so classed because they exist independent of conventional classroom-based institutions. Type One is represented by institutions such as International Correspondence Schools which was described earlier. Type Two is represented by institutions such as:

Allama Iqbal Open University, Pakistan
Athabasca University, Canada
Open University, Great Britain
Everyman's University, Israel
Fernuniversitat, Germany
Free University of Iran
Sri Lanka Institute of Distance Education
Universidad Estatal a Distancia, Costa Rica
Universidad Nacional Abierta, Venezuela
Universidad Nacional de Educacion a Distancia, Spain
(Kaye, 1981, P.15)

These institutions have many things in common. They are (for the most part) national universities rather than regional

universities. Their student enrolments are generally quite large. Harry (1990) notes for example that in 1987 there were approximately 120,000 students enrolled in Britain's Open University. In 1986 the Allama Iqbal Open University of Pakistan had an enrolment of nearly 65,000 (Satyanarayana and Koul, 1990). Another but more important common feature, however, is that they have emerged not as products of the conventional post-secondary educational system, but as independent entities. They have, in fact, taken the industrial model for educational design and delivery to its conceptual pinnacle.

Autonomous distance teaching universities spend tremendous amounts of time and money in course design and production. Coldeway (1988), for example, notes that Athabasca University allocates, on average, three months to produce a course design "blueprint", at least a month to review the design and potentially an additional month to revise the "blueprint". In total Athabasca spends a year or more developing a course at an approximate average cost of $125,000. In addition to this it spends approximately $1,000,000 per year employing 400 or so tutors who provide academic counselling to its 16,000 part-time learners (Coldeway, 1988).

Dual Mode Institutions

The mixed or dual mode institutions, of which Keegan (1986) says there are three types, rely on the academic resources and

infrastructure of the conventional post-secondary system to offer distance mode courses. Type one is essentially a North American model and is characterized by conventional universities which also offer distance mode courses. Type two are common in the former socialist republics of central and eastern Europe. Type three is an Australian model.

In the conventional system course development costs are the costs associated with employing faculty. Faculty are hired and expected to teach courses which, by virtue of their academic training, they are assumed to know how to teach. Dual mode institutions, then, are institutions which employ faculty to teach both in the classroom and to students who study at a distance. The variety of ways that this is done is reflected in the three types of institutions that Keegan (1986) says operate in this fashion.

The first type of dual mode institution is the type which is familiar to most Canadians and Americans. In Keegan's words, they are the "independent study divisions of a college or university" (p.148). They are likely the best known because they are the most numerous. There are many Canadian examples of this sort of institution. The following list represents most of them.

British Columbia:	Simon Fraser University
	University of British Columbia
	University of Victoria
Alberta:	Lakeland College

	University of Alberta
	University of Calgary
Saskatchewan:	University of Saskatchewan
Manitoba:	Brandon University
	University of Manitoba
Ontario:	Carleton University
	Guelph University
	Laurentian University
	University of Ottawa
	University of Western Ontario
	Wilfred Laurier University
Quebec:	University of Montreal
	University du Quebec
New Brunswick:	University of Moncton
	University of New Brunswick
Nova Scotia:	Dalhousie University
	Mount Saint Vincent
	Technical University of Nova Scotia
Newfoundland:	Memorial University of Newfoundland

For most of these institutions, distance education was a natural outgrowth of their activities in extension education which, as Swanson and Claar (cited in Lertpradist, 1990) point out, originates from the egalitarian conviction that the advantages of university education should be available to ordinary people.

Memorial University, for example, became heavily involved in educational television as a means of supporting courses that were being offered through extension (Rothe, 1986). And while Memorial's use of educational television eventually declined, it did retain a division of Educational Technology to provide media expertise to university faculty and extension workers who needed to convert their courses into courses usable for a distance mode. Like autonomous institutions such as Athabasca, Memorial and other dual mode institutions use academic tutors and a broad range of educational media to deliver their courses. However, their use of systematic instructional design in the course development process tends to be very limited. Dual mode institutions are inclined to adapt existing courses for the distance mode and for this they have been both commended and criticized (Knapper, 1985).

The difference between autonomous and dual mode institutions ultimately comes down to the issue of quality. For industrialized countries, particularly in the countries of western Europe, the United States and the "Pacific Rim" countries, the concept of quality and quality control in manufacturing has long been promoted as the major determinant of corporate success and longevity. The contest between North American and Japanese auto makers, for example, hinges largely on the quality of the product they each produce.

In conventional education there has traditionally been minimal emphasis on the use of quality control mechanisms in

course design. The quality of a given course, barring persistent complaints from students, is nearly always assumed to be high. The quality of distance mode courses which are delivered through the extension departments of conventional universities (i.e., dual mode institutions) is measured in a similar fashion. Autonomous distance teaching institutions, however, place a much greater emphasis on systematic instructional design, with its attendant emphasis on formative and summative evaluation, to ensure quality is built in. These institutions operate from the belief that quality distance mode courses are the product of an iterative and systematic design process and that such courses represent the combined efforts of course design teams consisting of (at a minimum) instructional designers/educational technologists and content experts/academics. Thus, when a representative of an autonomous distance teaching institution makes a comment like the following, the message that is conveyed is that distance mode courses offered by instructional design oriented autonomous distance teaching institutions are inherently better:

> We have not exactly mailed out lecture notes with a series of textbooks and said, go for it. We've done a lot of work that a lot of people in Instructional Design would call meaningful work (Coldeway, 1988).

The other two types of dual mode institutions which Keegan

describes (i.e., distance education departments in the former socialist republics of Central and Eastern Europe and the Australian integrated mode (New England model)) vary from the first mostly in terms of the way the institutions themselves have been structured to deal with distance students. In the socialist model,

> [l]earning materials for use throughout the nation are developed by course teams of professors....The materials are then distributed to the institutions which are going to enrol and teach students in that particular discipline. On enrolment students are allocated both to the institution from which they will get their degree (which may be far away) and to a consultation centre at an institution near to their home or work. Study commences with a residential seminar on campus after which students study at home from the learning materials provided (Keegan, 1986, p.152).

This model of distance mode instruction is minimalist compared to western systems because, as Keegan (1986) points out, 'corresponding' plays a very small role. Learners are enrolled in a parent institution and assigned to a satellite institution at which, following regular periods of self study, they attend classes and take examinations (Keegan, 1986). It might therefore be better described as conventional education with very few classes.

The last type of dual mode system which Keegan describes is the Australian integrated mode (New England Model). Distance education, external studies as it is known in Australia, has had a long history. Dating from as early as 1910 (Keegan, 1986), Australian distance education has evolved as a response to the needs of a sparsely populated continent. Emerging from the practice of using communications technology to deliver conventional style instruction, the Australian distance education approach resembles the dual mode approach common in North America.

In the New England model, two of most important factors to precipitate the rise of distance education were the advent of modern communication technologies and the need to provide professional upgrading to high school teachers living and working in remote areas. Improved communications technology, especially, was important as a mitigating factor for the logistical problems associated with teaching small numbers of people scattered over a very wide area (Smith, 1983). Australian integrated mode universities, while they resemble the dual mode institutions found in Canada and the United States in terms of their course development and delivery approach, tend to place a greater emphasis on having some face-to-face contact between the learner and the instructor or tutor (Smith, 1983). Some schools such as the University of New England (whence came the name "the New England model") have compulsory study schools for its distance mode students. While this is a common practice for Australian

integrated mode institutions (Thornton, 1990) it is not a common requirement for integrated mode institutions in North America. Smith (1983) notes that in New England model universities " [the] academic staff are responsible for the total teaching/learning process of writing courses, teaching them through a combination of independent study materials and face-to-face tuition and assessing the students by way of assignments and formal examinations" (p.199).

The criticisms that autonomous distance teaching institutions have of institutions which do not share their approach toward course development, and thereby quality control, therefore applies to all three types of dual mode institutions that Keegan (1986) describes. As a consequence, the rhetoric surrounding the measurement of the merit and worth of a given type distance education can be rather perplexing not only for the consumer of distance mode courses but for the researcher and practitioner as well.

Evaluation Within Distance Education

Evaluation has traditionally been given a higher priority in distance education systems than in conventional educational systems (Alvarado, D'Agostino, Bolanos, 1991). This is particularly true for autonomous distance education institutions. Conventional educational systems, including those which operate in a dual mode (i.e., conventional education plus distance

education) tend to look at the economics of distance education in different ways. Conventional universities tend to economize on existing infrastructure to offer a service rather than construct courses and offer a service at a level of quality that would be economical only when offered through a large client/student-base -- i.e., through the economies of scale (Stubbs, 1985) .

The Economics of Conventional Education

Conventional universities, using a conventional classroom approach, sometimes have faculty teach in large lecture theatres so that as many students as possible can hear the live lectures. And when lecture theatres are not large enough some universities even resort to using closed-circuit television to broadcast faculty lectures to classrooms elsewhere on campus. As classrooms get larger, the opportunity for direct and immediate communication between the student and teacher diminishes. When students must watch and listen to faculty from television monitors in remote classrooms, the opportunity for immediate communication has clearly disappeared. At this point, it really does not matter whether the televised lecture is live or taped. Thus, the lecture (justifiably or not) has been given a supremely important position in the learning experience at conventional universities.

The Economics of Distance Education

Converting a conventional classroom-style course for use in a distance mode typically involves the transcription of lecture notes (possibly supplemented by audio or video tapes of some of the lectures). It is, for most dual mode institutions, a conceptually simple, procedural task. Moreover, from the perspective of the student, there is probably little qualitative difference between this type of distance mode course and its conventional 'lecture theatre' counterpart other than the fact that in a distance mode, the learner has greater flexibility in terms of instructional time.

Flexibility is a word that is often used to promote distance mode courses. Unfortunately, flexibility is not necessarily a positive attribute. This is especially true if the flexibility is more a consequence of the course delivery approach than an intentionally designed feature of the course. Most students are not prepared for the self-study nature of distance mode courses. Additionally, the needs of distance mode students are very often different than those of students in conventional education systems (Stubbs, 1985). Ignoring these differences can lead to frustration, failure and dropout. Indeed, there is evidence which suggests that high drop-out rates, some as high as 70 percent, are common in distance education (Woodley and Parlette, cited in Garrison, 1987; Coldeway, 1988).

Another consideration about the flexibility concept is that

the average student in a distance education course tends to be a little different than the average student in a conventional classroom setting. Indeed, since distance education has evolved as a more egalitarian approach to education, it has attracted students who would ordinarily be inhibited from attending conventional institutions because of barriers such as geographical isolation, full-time employment, financial constraints, family commitments or physical disability. In fact, autonomous distance teaching universities have targeted these populations and have been committed to offering a high quality education, with the removal of the barriers that have traditionally kept many from participating in post-secondary education. This is particularly true of those autonomous institutions which use the term "Open" as a part of their institutional name. Dual mode institutions, while they may have similar goals, are tied to supporting the infrastructure of the conventional university system, a linkage can work as a handicap.

At Athabasca University the typical student is a 35 year old woman who works full time and has some coursework completed toward a degree (Coldeway, 1988). At Télé-Université, the only other autonomous distance teaching university in Canada, enrolment is balanced between men and women (Stahmer and Helm, 1987), most are part-time students and 40 percent fall into the 30 to 40 age range category (Guillemet, Bedard and Landry, 1986).

Learner-centred Versus Institution-centred Education

Autonomous distance teaching institutions tend to be highly learner-oriented. Télé-Université, for example, provides an extensive range of student support services including regularly scheduled, tutor-led discussion groups, where a group of students in a particular community makes it practical. It offers regularly scheduled telephone coaching and teleconferencing, especially for students in remote areas; it offers small, unmoderated study groups where the nature of the course and the geographic proximity of the students make it possible; and it offers individual tutoring support and combined methods when necessary. Also, courses are subject to ongoing evaluation through questionnaires sent to students and through reports from tutors (Lamy and Henri, 1983). Print materials developed by Télé-Université course teams are evaluated every three years (Guillemet, Bedard, and Landry, 1986).

Dual mode institutions, by contrast, tend to be institution oriented; that is, the mode of operation is centred more in the value to the institution, rather than to the learner (Stubbs, 1985). Thus, they tend not to place as much emphasis on student support or course evaluation. Such institutions are inherently constrained by the philosophy and practice of the conventional university system of which they are a part.

Conventional universities usually do very little in the way of quality control for course development because courses are

typically developed by faculty who, working without the assistance of instructional designers and/or educational technologists, focus mostly on the content of what they want to teach. The traditional emphasis on individual academic autonomy at conventional universities runs counter to the belief held by professional distance educators and researchers that the development of quality distance education courses requires the expertise of people knowledgeable about learning theory and instructional design in addition to content experts (i.e, faculty) (Dillon and Gunawardena, 1992).

The pressure on dual mode institutions to adhere to the principle of academic autonomy and to sustain the costly physical plant required to operate a campus-based educational system is enormous compared to the pressure to apply higher standards and budgets for the development of distance mode courses. This means that the treatment accorded the development of distance mode courses is much like that accorded the development of conventional courses.

Equalizing Autonomous, Dual Mode and Conventional Systems

The argument being presented here is that classroom-based courses are necessarily weakened when they are converted for distance mode use and that additional measures (at a minimum, a systematic and professionally executed evaluation) need to be implemented to ensure that distance mode courses offered through

dual mode institutions are at least as effective as their conventional mode counterparts are presumed to be.

While there are many obstacles which could thwart efforts at dual mode institutions from implementing regular and systematic educational program evaluation for distance mode courses, none should be so insurmountable as to prevent evaluations from being done. Until recently, one of the major obstacles that prevented dual mode institutions from investing in quality control systems was, ironically, the fact that the conventional "orientation" of the decision makers accorded distance education a second-class status (Stubbs, 1985). But, as Evans and Nation (1990) point out, distance education is no longer a marginal activity. Its feasibility and effectiveness, although not always assured, has been proven. Academic arrogance for conventional style education is no longer justifiable. The only apologies that distance educators should make now are for failing to design and test workable program evaluation models for the increasing variety of subject matter and approaches that are being offered through the distance mode.

Program Evaluation

The Oxford Dictionary (1989) defines evaluation as follows: "find or state the number or amount of; appraise, assess". The root word is value. Oxford defines value as "worth or desirability or utility, or qualities on which these depend".

Based on these definitions then, program evaluation should be the process of quantifying, appraising or assessing a program for the purpose of determining its worth or desirability or utility.

Thus, one might expect to find fairly consistent definitions for program evaluation in literature on the subject. However, such is not the case. Rutman (1980) says, for example, that "[t]he term 'program evaluation' does not have a standardized and commonly accepted meaning. Rather, there are widely different interpretations of the term" (p.17). Guba and Lincoln (1989) clarify this seemingly perplexing situation when they say that to define evaluation is to imply that there is a "right" way to evaluate.

Evaluation is a construct, "a set of theoretical and practical activities without a widely accepted paradigm" (Glass & Ellett - cited in Schuemer, 1991, p.5). Berk and Rossi (1991) note that it is a concept which derives from a common-sense idea about the value of judging and the judging of value. Thus, in a general sense it is a process with many acceptable delineations and a set of kindred purposes which include "judging decision alternatives" (Committee on Evaluation, 1971,; Stufflebeam et al, 1971), judging merit, value and worth (Borg & Gall, 1989), reducing uncertainties, improving effectiveness and making decisions (Patton, 1982), "the assessment of program impact and the analysis of program benefits relative to their costs" (Berk & Rossi, 1991, p.8), and improving decisions (Thompson, 1975).

Program Evaluation in Historical Perspective

The need or desire to know the value of something is undoubtably as old as civilization itself. Indeed, evaluation (i.e, judging) is an inherent part of civilization. Once value is established, regardless of how temporal that value may be, decisions related to that value can be validated. As a concept, therefore, evaluation is as old as the process of decision-making.

Nevertheless, formal evaluations such as that connoted by the term program evaluation are a relatively recent phenomenon. Even in the field of education, an area where evaluation would seem to be a supremely important principle, formal evaluation can be traced only to the latter half of the 1800s. Worthen & Sanders (1987) note that prior to 1837, decisions in education were either (or both) politically or religiously based.

A series of events during the nineteenth century in North America aroused public concern about and focused attention on education. These events included the increasing levels of child poverty and crime in the growing American and Canadian cities, the massive waves of non-English speaking immigrants flowing into rural and urban areas of the United States and Canada, and the subsequent xenophobia. Education came to be seen as an essential or necessary tool for the assimilation and socialization of a rapidly expanding North American population. Tyack (1976) notes, for example, that nineteenth century Americans "had enormous

faith in the power of schooling to transform all kinds of people - even "enemies" - into citizens" (p.365). Fox (1991), in fact, traces this faith in the power of education to the American and French revolutions. He says that since that revolutionary period, "education has been used to induce knowledge in learners on a mass scale as a necessary social prerequisite to political democracy" (pp.218-219).

Nineteenth Century Quest For Improvement

Horace Mann.

As already noted, the nineteenth century was the blossoming era of the Industrial Revolution. It was the era of frontiers and individualism, of institutional refinement and the democratic ideal, of citizenship and leadership. And thanks to Horace Mann, one of the century's most progressive social activists, it was the era that saw both legislation and public commitment for improved standards in public education.

In 1837 Mann, a successful and well-established Massachusetts lawyer, was appointed Secretary to the State's Board of Education and over the following twelve years he shaped the course of education not only for Massachusetts, but for the whole of the United States and much of Canada (Downs, 1974). In each of the twelve years of his tenure (1838-1850) Mann produced an Annual Report on current educational concerns in the State.

These reports, which are today still available on the shelves of university libraries throughout North America, were the result of what were then seen as empirically supported research investigations (Worthen and Sanders, 1991). Significantly, however, these reports also stand as harbingers of modern educational and program evaluation.

Egerton Ryerson.

Egerton Ryerson, Canada's nineteenth century counterpart to Horace Mann was a similarly imbued social activist. And although his strong Methodist upbringing coloured his view of educational reform, his methods and ultimate objectives were similar to Mann's.

Ryerson, like all nineteenth century educational reformers, in his quest for educational improvement had to grapple with the notion of quality. For Ryerson, this was a question of how to distinguish the good from the bad in the conventional wisdom about educational practice. After all, it is one thing to make assertions about the quality of a program, but it is quite another to find credible and valid support for these assertions. Ryerson, and others of his era, derived this validation from his own moral standing and intellectual and oratorical prowess amongst his cohort of political influentials.

In a prefatory note to Ontario Provincial Secretary concerning his 1846 Report on a System of Public Elementary

Instruction Ryerson made the following statement:

> I have 'borrowed from all [i.e., all sources documented in the report] whatever' appeared to me to be 'good', and have endeavoured to 'perfect', by adapting it to our condition, 'whatever I have appropriated'" (cited in Fiorino, 1978, p.59). Thus, the criterion against which he made his value judgements was that which appeared good to him. As Fiorino (1978) points out, Ryerson's concept of 'good' (in this case as it pertains to his conception of education) was related to "his view of the good of man, that is, in relation to the moral dimension of his thought". (p.59)

The Evolution of Evaluation Methods

While the efforts of individuals like Horace Mann and Egerton Ryerson may reflect the era in which it became important to try to collect empirical data to support certain assertions about education (Worthen and Sanders, 1987), their methods would not be classified as scientific by today's standards. On the other hand, theirs was an epistemology which derived from the social, cultural, religious and political bias of the time. At the time, and indeed retrospectively, their ideas about improvement were progressive. What has survived from their efforts, however, was their agenda for educational improvement.

From the time of Mann and Ryerson, several generations of

progressive thinkers have had significant directional influence on the path that leads up to the modern practice of educational and program evaluation. The testing movement of the early 1900s, for example, set a trend toward the use of norm and criterion referenced tests to quantitatively measure human abilities (Worthen and Sanders, 1987). In fact, testing gained wide acceptance in the United States military during World War I and in private industry during the post war era (Worthen and Sanders, 1991) The next major milestone began in 1932 with the work Ralph Tyler, a man whose ideas about and approach to evaluation persist to this day.

Ralph Tyler, alternately called the father of educational evaluation and the father of behavioural objectives, gained prominence during the Great Depression. He says that at that time "[p]eople were worried about their material losses and blamed much of it on the banks, the government, and the schools" (Nowakowski, 1983, p.24). The depression had caused a dramatic increase in school enrolment, primarily because employment was virtually unattainable and young people saw few other options. Traditionally, the curriculum for most high school programs had been oriented toward college entrance, yet an increasing percentage of the students were not interested in going to college. The state of Ohio responded to this problem by proposing an eight year long experiment that would develop and implement new educational programs to serve the needs of a more diverse group of students. A stipulation of the plan, however, was that

the performance of the students in these new programs had to be evaluated. Tyler was selected to direct the evaluation because he was known for evaluating schools based not on tests, but on objectives (Nowakowski, 1983).

A focus on objectives.

Ralph Tyler's famous Eight Year Study of the progressive new curriculum devised by the State of Ohio became a landmark in the history of educational evaluation. Tyler pointed out that the goal of instruction is to achieve certain objectives. He argued, therefore, that if the objectives are stated in behavioural terms they can then be used to evaluate the effectiveness of the instruction because behaviour is both observable and measurable (Reiser, 1987). Thus began a transition from a program evaluation model which relied on an appraisal of students using tests which served little other purpose than to provide grades, to a model which appraised students' mastery of the objectives that program designers had outlined before the program began (Popham, 1988; Tyler, 1980).

Industrialization, World War II, the post-war era.

The Industrial Revolution as an intellectual movement was tested for its authenticity by World War II. The marshalling both of knowledge and the creators of knowledge for the war effort had

a profound effect on the direction of civilization. New emphasis had been placed on the quality of the inputs that were being fed into the great industrial machine. Winning the war was contingent on having the best technologies, strategies, and people, and finding the best became a new twist on the industrial ethic. Nowhere was this more true than in the areas of psychology and education. Indeed, many of America's best and most progressive academic minds went to work with the United States Department of Defence (Gagne, 1989; Saettler,1990). Ralph Tyler, for example, became the Director of the Examinations Staff to develop educational testing for the armed forces (Nowakowski, 1983).

The Allied Countries emerged from the war with the knowledge that vigilance was needed in education and the development of human capital if future world peace and stability were to remain viable expectations. The United States, in particular, had positioned itself in a leadership role -- being the best, and not being outdone, became a part of national culture. Hence, when the Soviet Union launched its space satellite, Sputnik, in 1957, before the United States had even established a space program, American national pride was bruised and the public school system was singled out as one of the major causes. As a consequence, considerable new Federal money was allocated for the development of new approaches to curriculum development for mathematics and science (Popham, 1988).

Fortunately for the United States, the infusion of new money into education was timely because the work of American

psychologist Benjamin Bloom and his colleagues had been published just one year before the launch of Sputnik. Although the launch of Sputnik received much more attention, Bloom's *Taxonomy of Educational Objectives* (1956), while not immediately recognized for its significance (Reiser, 1987), became a landmark for research and practice in education and psychology and a standard for students in the fields of learning and instruction and evaluation. Hence, many of the post-Sputnik educational initiatives relied on foundations provided by individuals like Tyler, Bloom and associates, M.D. Engelhart, E.J. Furst, W.H. Hill, and D.R. Krathwohl, and many others who were part of what is sometimes called the objectives-movement.

The late 1950s and the early to mid-1960s especially were formative years for program and educational evaluation. B.F. Skinner's (1954, 1958) ideas about applied learning theory and programmed instruction, Robert Mager's work during 1960s on the use and preparation of objectives for instruction (Mager, 1962), Robert Gagne's (1962) work on the classification of instructional objectives, and their instructional outcomes, and numerous other publications by these and other researchers working in that era (Gagne, 1987) evolved with the work of Ralph Tyler and Benjamin Bloom to stimulate the growth of the instructional systems development movement during the late 1960s and early 1970s. Instructional systems development became the crowning product of much of the post-war research on learning and instructional theory and represented a new school of thought that promoted

multi-stage, iterative models for designing and developing instruction that was objectives-based, and both formatively and summatively evaluated (Romizowski, 1981; Reigeluth, 1983; Gagne, Wager & Briggs, 1988).

Educational program evaluation, while it was heavily promoted by the instructional systems movement, has evolved as an establishment, "with its own organizations, publications, institutions, and ways of behaving" (House, 1986, p.5). Indeed, research in the field of evaluation, its epistemology, ideology and methodology have been at the heart of developments in instructional systems development.

Throughout the 1960s and 1970s the practice and theory of evaluation evolved and was refined along new dimensions. Over that period researchers began to question the assumptions inherent in the traditional positivistic (quantitative) approach to research and evaluation, thanks partly to Thomas Kuhn who, in his 1962 book *The Structure of Scientific Revolutions*, challenged researchers in all the fields of "natural science" to recognize the impact their guiding paradigms were having on their ability to make research-based discoveries.The positivistic approach was also challenged by the growing acceptance of the validity of qualitative research and evaluation methods and approaches that emerged out of fields like anthropology, sociology, history and economics (Borg and Gall, 1989); These factors, combined with the fact that until the late 1960s there was a dearth of practical and tested models, strategies and plans for evaluators to follow

(Worthen & Sanders, 1973), led to the emergence during the 1970s early 1980s of a significant body of new and practical models and approaches to program evaluation, and a new concern for professional standards of practice (Patton, 1982). Indeed, Michael Quinn Patton's book *Practical Evaluation*, published in 1982, reflects the modern age of evaluation. He says,

> [a] practical evaluation is doable and applicable. It is doable in that the design is feasible and can be implemented within the financial, time, and political constraints of a particular situation. It is applicable in that the evaluation findings can be used, i.e., appropriately and relevantly applied, by decision makers and information users (p.296).

The important thing to note about this comment is that it focuses on concepts like evaluation design, feasibility, constraints, and audiences. These are the sorts of characteristics which typify modern evaluation models.

A Taxonomy of Evaluation Models

There are numerous evaluation models for professional evaluators to choose from and, as House (1978) says, one way to understand the route that modern evaluation has taken is to compare the models. He says that by comparing the models we see

how logically similar they are to one another. He says further that the assumptions implicit in all the major models are essentially variations in the assumptions that are associated with the conceptions of a liberal democracy, itself an outgrowth of "an attempt to rationalize and justify a market society" (p.5). In this respect, evaluation plays an instrumental role in the industrial transformation of education (House, 1986).

To facilitate the comparison of the major evaluation models House (1978) created a taxonomy which compares each model on principal proponents, major audiences, principal evaluation measures, methodology, outcome and typical questions. His ordering of the models (Systems Analysis, Behavioral Objectives, Decision Making, Goal Free, Art Criticism, Accreditation, Adversary, and Transaction) is particularly useful in respect to this study because it can be used as an aid in describing the rationale for selecting the evaluation model that became the basis for the study.

House (1978) presented his taxonomy in a table format (see Figure 1). He states:

> [i]n the taxonomy the models are related to one another in a systematic way. Generally the more one progresses down the column of major audiences, the more democratic or less elitist the audience becomes. The more one moves down the consensus column, the less consensus is assumed on goals and other elements. The more one moves down the methodology

column the more subjective and the less objective the research methodology becomes. The more one moves down the outcomes column, the less overall concern becomes social efficiency and the more it becomes personal understanding (p.5).

A TAXONOMY OF MAJOR EVALUATION MODELS

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

Figure 1 A taxonomy of major evaluation models (House, 1978, p. 12).

Thus, the transaction model, the last evaluation model in the first column is described as being the most democratic, the one which assumes the least consensus on goals and other

elements, the one which is the most subjective and the one which adopts personal understanding as the primary concern. And from this model was borne the evaluation approach that was used for this study.

The Major Evaluation Models

The following is an assessment of each of the models listed in Figure 1. It addresses each model in the order they are listed in the taxonomy and describes each in terms of its salient advantages and disadvantages.

Systems Analysis

The desired outcome of a systems analysis evaluation model is an objective statement of the effectiveness and efficiency of the program it is used to evaluate (Patton, 1982; House, 1978). House (1980) says that the systems approach is characterized by the quantitative analysis of data such as test scores, the objective of which is to find statistical relationships between output measures and the programs being evaluated. However, a more thorough understanding of the systems analysis approach can be achieved by looking at its historical origins.

The systems approach, a derivative of systems theory, emerged in the era between the two World Wars as the foundation for scientific research methodology. It was a theoretical

principle which guided the practice of research and was essentially a joint quality control and problem solving heuristic for scientists. It encouraged them to think of their experiments both in terms of the elements which comprised them and the interaction between the elements.

The benefit of using the systems approach was that it operationalized a highly controlled and, in principle, an inherently simple technique for problem solving. In 1945 the famous American mathematician G. Polya introduced an approach to mathematical problem solving which exemplifies the systems approach. Polya said that to solve a mathematical problem,

1. You must first understand the problem.
2. You must find the connection between data and the unknown and obtain a plan of the solution.
3. Carry out the plan.
4. Examine the solution obtained.

(cited in Romiszowski, 1988, p.31).

Polya's approach was essentially a derivative of the Scientific Method. Indeed, as Saettler (1968) notes, the systems approach implies the use the scientific method.

Rossi and Freeman (1993), authors of the widely regarded standard text on the use of systems analysis for program evaluation, state that their book,

> is rooted in an approach that has aspired to be effective in improving the quality of our physical and social environments and enhancing the chances of our individual and collective survival: the application of scientific procedures to social problems. If the term *evaluation research* is a relatively recent invention, the activities that we will consider under this rubric are not. They can be traced to the very beginnings of modern science in the 1600s. As Cronbach and his colleagues (1980) point out, Thomas Hobbes and his contemporaries four centuries ago were concerned with devising numerical measures to assess social conditions and to identify the causes of mortality, morbidity, and social disorganization (p.4).

In his taxonomy of major evaluation models, House (1978) cites Alice M. Rivlin as the major proponent of the approach. And, at the time of the publication of that taxonomy in 1978, Rivlin was considered the leading advocate for the approach. However, if House were to revise his taxonomy, Peter H. Rossi and Howard E. Freeman (deceased) would certainly have to be placed in the same category.

At the time of the publication of her book *Systematic Thinking for Social Action*, Rivlin was an economist and senior fellow of the Brookings Institution. From about 1965 until the late 1970s systems analysis had been the dominant evaluation perspective for the United States Department of Health, Education

and Welfare (House, 1978; Rivlin, 1971). Rivlin's book was an effort at evaluating the overall effectiveness of that approach (Rivlin, 1971). Rivlin (1971) argues that to be assured of doing a good job at running a government program, administrators should adhere to the following steps:

> (1) Define the objectives of the organization as clearly as possible;
> (2) find out what the money was spent for and what was being accomplished;
> (3) define alternative policies for the future and collect as much information as possible about what each would cost and what it would do;
> (4) set up a systematic procedure for bringing relevant information together at the time decisions were to be made (p.3).

Sungalia (1980) says that principal concern of the systems analysis approach is to ensure that what is invested in an educational program (she cites inputs like professional expertise and time, facilities and materials, etc.) is cost-beneficial. She says that systems analysis evaluators ask questions like, "Is the program producing sufficient educational benefits for the costs incurred?" or "Is the program producing a particular benefit, more or less expensively per unit of program outcome, than other programs designed to achieve the same objective?" (p.2).

In summary, the systems analysis approach to evaluation leans toward highly quantitative and objective analyses of the relationship between program inputs and outputs. As House (1980) puts it, the objective is to establish cause and effect relationships between the programs and their outcomes. It attempts to assess such things as cost benefit and efficiency and provides its major audiences, namely managers and economists, with data from which to guide decisions relating to management, planning, policy development, and for fiscal purposes (House, 1980; Rossi and Freeman, 1993).

Advantages.

When the dominant concern on the part of those who fund social programs is cost effectiveness, and when the preference is for program outcomes that are stated in quantitative terms, systems analysis should be considered as a potentially desirable evaluation model. Data from systems analysis evaluations is highly conducive to statistical manipulation and can be adapted to a wide variety of graphic presentation formats. It is therefore relatively easy to reduce large quantities of data into presentations that quickly disseminate evaluation results.

An additional consideration is that because systems analysis emphasizes the use of quantitative measures, often through the use of survey research procedures, and because of the advent of sophisticated computer-based statistical software, evaluators

using this approach are able to analytically examine large numbers of variables (Rossi & Freeman, 1993). This is advantageous for several reasons. It makes it feasible to evaluate very large programs and to use large population samples within time and economic constraints that would likely prove too onerous for more subjective, ethnographic evaluation approaches. Moreover, since statistical validity and reliability studies which have many uncontrolled variables are impacted by sample size (Borg and Gall, 1989) systems analysis evaluations will require large sample sizes. Another reason is that if small correlational relationships between program factors are anticipated, it is desirable to use large samples (Borg and Gall, 1989).

Disadvantages.

The major audiences for the systems analysis approach are economists and managers (House, 1978; 1980) which makes it elitist in terms of its presuppositions about things like the types of data that ought to be collected, how it should be collected, interpreted and used (Robertson, 1982). In other words, input about what standards and criteria should be used in the evaluation, and who should play a role in defining them is limited to a select few because, as House (1980) says, "a consensus on goals is assumed" (p.25) (see also Webber, 1987, p 4.). The problem with assuming a consensus on goals is that the

evaluation will produce biased and thereby distorted findings about program effectiveness.

Another disadvantage of the systems analysis approach is that because it relies on statistical analysis techniques to interpret data and because validity in statistical analysis is related to sample size, it is not well suited to evaluating programs in which there are few participants. Indeed, it is not well suited to evaluating any program which by design or operation does not generate sufficient quantities of measurable quantitative data to provide statistically valid results.

A further problem with systems analysis is that because it relies on quantitative measures, the number of factors that can be investigated is constrained. Indeed, certain factors do not lend well to quantitative measurement (Webber, 1987). Moreover, it is unable to measure impacts (negative or positive) which were not intended. Unintended program impacts can provide significant information to program developers and policy makers alike and if they are not identified or explored, judgements about the program must be tempered by the knowledge that findings may be only a narrow reflection of its full value (Scriven, 1986).

Further, systems analysis evaluations are not well suited to programs which are piloted for the purposes of providing multiple audience groups with feedback about the program. The type of evaluation finding that is desired by program designers and administrators is not the same as that which would be desired by policy makers or fiscal managers. Moreover, since systems

analysis evaluations are designed to provide policy makers (e.g., economists) and fiscal managers with information about which to make decisions regarding continued program funding, and since program designers and managers desire information that will enable them and other similarly interested audiences to make program improvements, systems analysis evaluations are not well suited to the purposes of the latter audiences. Thus, strong consideration needs to be given to the reason for the evaluation so that an appropriate model may be selected (Borg and Gall, p.745).

Behavioral Objectives

The focus of the behavioral objectives model, alternatively known as the goal-based model, is the objectives or goals which the program is intended to achieve. According to this model, program success can be measured by an assessment of the discrepancy between the stated objectives and program outcomes (House, 1980). Developed by Ralph W. Tyler, the behavioral objectives model has dominated thinking about evaluation, particularly educational evaluation since the 1930s (Worthen and Sanders, 1987).

According to Tyler (1949), an objective has four elements: "a) definition of purpose, b) experiences needed to achieve the purpose, c) the organization of experiences and d) method and standard for determining attainment" (p.1). One of the notable

aspects of this definition is that it emphasizes experiences. The behavioral objectives model measures participant experiences, stated as behavioral objectives.

Tyler's behavioral objectives approach was concerned primarily that program objectives and thereby the evaluation objectives were stated in behavioral terms. The belief was that if the desired behaviour could be stated or described, it could be measured with some sort of achievement test. In his refinement of objectives-oriented evaluation Robert Mager (1962) said that it is not simply enough to state the objectives, they must also include the desired attainment levels and the criteria by which these attainment levels can be judged (Worthen and Sanders, 1987).

Other evaluators have re ined or modified Tyler's objectives-based approach, using it as the general model, to better serve, as they see it, the evaluation function. One such variation of the Tyler approach is discrepancy evaluation developed by Malcom Provost. Like Tyler's approach, Provost's approach assesses a discrepancy between program objectives and program outcomes. However, Provost's discrepancy approach provides both formative and summative information about the program whereas Tyler's approach provides mostly summative information (Provost, 1973; Stufflebeam & Shinkfield, 1985). In other words, Provost's discrepancy approach provides findings that say not only whether the program was effective, it also provides information about ways to improve the program. Tyler's

approach is concerned more with program effectiveness.

Another variation of the basic behavioral objectives model is cost-benefit evaluation which, in function at least, is similar to the systems analysis model of evaluation. Evaluators using the cost-benefit model attempt to assess program objectives in relation to the cost needed to achieve them (Borg & Gall, 1989).

Whatever the variation, the behavioral objectives model focuses on initial program objectives. They are, in effect, the standards against which evaluators using this model measure or judge the program.

Advantages.

The behavioral objectives model is advantageous to the evaluator because it simplifies the task of developing suitable instruments for measuring the learner's achievement of program objectives (Borg & Gall, 1989). The problem of developing standards with which to measure the program is virtually eliminated because they already exist in the form of program objectives or goals. Moreover, since standards already exist, the step of creating instruments with which to measure them is reduced to a largely procedural task.

Worthen and Sanders (1987) note that the greatest strength of the objectives-based evaluation model is its inherent simplicity -- "It is easily understood, easy to follow and

implement, and produces information that educators generally agree is relevant to their mission" (p.72). Moreover, because of its inherent simplicity, the objectives-based model (particularly Tyler's approach) can be used not only by professional evaluators, but by individuals with a limited background in program evaluation as well (Stufflebeam & Shinkfield, 1985).

There is also benefit to the use of the behavioral objectives model by virtue of the fact that it is widely known and used, probably more than any other model. There have been ample evaluations of educational programs produced using this model to use to gauge the merit of using the model for a given program. In addition, there are a sufficient number of approaches to select from to provide the level of formative or summative information necessary to satisfy a range of audiences.

Disadvantages.

Critics of the behavioral objectives model have asserted that it

> (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 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 program value not reflected in its own objectives, and (9) promotes a linear, inflexible approach to evaluation (Worthen & Sanders, 1987, p.73).

Another major criticism of the behavioral objectives model is that it does not adequately consider differences in the experiences of learners, abilities or interests (Brookfield, 1986). Indeed, the same criticism is levied against all of those models that House (1978) describes as epistemologically objectivist, namely, systems analysis, behavioral objectives, decision making and goal free. These models, he argues, are concerned more with social efficiency than with personal understanding.

Brookfield (1986) says that the behavioral objectives model, and presumably other similarly objectivist models, are incompatible with the "democratic ethos" of programs developed through an Adult Education ideology (pp. 267-268). The collection of evaluative data designed to inform an elite audience (e.g., managers or decision makers) is in opposition to Adult Education theory which advocates the participation, at all levels, of a broad group of the audience, including program participants.

Decision Making

The decision making model of evaluation derives its name from the way the model is structured and the way the findings are used. Its function is to provide decision makers such as administrators and managers with information they can use to make decision alternatives (House, 1978; Worthen and Sanders, 1973; Webber, 1987). Worthen and Sanders (1973) note that evaluation is sometimes defined as "the ascertainment of value and decision" (P.129), which provides a rational for a model that is structured according to the actual decisions to be made (House, 1980).

The decision making model is intended to enable managers to make critical decisions about how a program can be improved (i.e., quality control) or whether and to what extent a program should be continued (i.e., effectiveness), or both. Each of these decision types represents a different reason for enquiry and consequently, different methods of gathering evaluation data. The most widely known and used application of the decision making model is an approach developed evaluation theorist, Daniel Stufflebeam and is known by its acronym, CIPP (Context, Input, Process, Product). Each element of the CIPP approach represents a different type of evaluation, and each "is tied to a set of decisions that must be made in planning and operating a program" (Borg and Gall, 1989, p.767). Nevo (1986) says, in short, that the CIPP approach assesses the merits of a program's goals, the quality and extent to which the plans are carried out, and the

worth of its outcomes (p.19). Moreover, it is through the analysis and synthesis of the findings from each of these evaluation types that the evaluation approach is able to provide data on the quality control in, or the effectiveness of, the program being evaluated. Because the CIPP approach is the source of most of the available data about the decision making model, for the purposes of the following discussion, the decision making model will be represented entirely by the CIPP approach.

Advantages.

For evaluators who require an evaluation model that would not place them at arms length from the program being evaluated (Borg and Gall, 1989), the CIPP approach ought to be considered as a possible option.

Brookfield (1986) says that because evaluators using the CIPP approach must evaluate the initial policy decisions that give rise to the program, and because they must also evaluate the way the program operates, the approach is very useful for informing the concerns of program practitioners. Brookfield (1989) notes further that the approach allows for the acknowledgement of concerns for "the influence of institutional priorities, the impact of individual personalities, and the prevailing political climate" (p.270).

The CIPP approach is, by design, comprehensive (Borg and Gall, 1989), and has been used extensively in the evaluation of

education programs (Cross, 1992). Cross (1992) notes also that findings can include information concerning unexpected outcomes. Moreover, because of its comprehensiveness, evaluator recommendations or decisions are considered thoroughly informed.

Disadvantages.

Because the CIPP approach is structured by the decisions to be made, and because decision makers define the decisions from which the evaluation questions arise, the questions that are considered relevant will reflect the interests of a narrow group. While some evaluators may consider this an advantage because of the belief that it will increase the utilization of the evaluation findings (House, 1980), the disadvantage is that while the findings may reflect a thorough investigation, the evaluation is at risk of being biased from the start by focusing on the interests of an elite audience.

The most widely discussed concern about the CIPP approach is its cost. The comprehensiveness for which the approach is praised is a double edged sword because thorough investigations are expensive on money and time. Brookfield (1986) says that "[t]o conduct an evaluation of all stages of a program's development may consume more time and energy than that expended in actually executing the program" (p.270).

Goal Free

The goal free evaluation model was developed by Michael Scriven as an alternative to the goals-based models that dominated during the 1960s and 70s. Scriven's purpose in developing the model was to make available an approach to evaluation that reduced evaluator bias (Salasin, 1974). Completely contrary to the goals-based models, the goal free model purposely ignores the goals that are set for a program and attempts to discover all program effects irrespective of the intentions of the program developers (House, 1978; Stufflebeam & Shinkfield, 1985). The findings produced by this model therefore reflect not what the program developer had intended to achieve, but what was actually achieved (Salasin, 1974).

The major audience for the goal free model, which strongly parallels Ralph Nader's work in the field of consumerism (Stufflebeam & Webster, 1983), is the consumer (House, 1978: Salasin, 1974). The purpose of a consumer oriented evaluation study "is to judge the relative merits of alternative educational goods and services and, thereby, to help taxpayers and practitioners to make wise choices in their purchase of educational goods and services" (Stufflebeam & Webster, 1983, p.34).

Advantages.

One of the clearest advantages of the goal free model is that it produces findings about unintended effects which, as Scriven argues (see Salasin, 1974: Worthen and Sanders, 1987), the goal-based models are inherently inclined to miss. Since the consumer (i.e., potential program participants) is the intended beneficiary of human service programs it should be a desirable objective to demonstrate for that population the full effect and value of that program. Moreover, although the consumer is the intended audience, it should also be desirable for managers, administrators and program developers to know the true value the program can provide that audience.

Another strong advantage of the model is that it is very concerned with controlling the level of bias that may enter the evaluation findings. (Innes, 1982) points out that since evaluation data is collected without the bias controls that are standard in true experimental designs, evaluators must remain acutely aware of, and at every opportunity attempt to control, all sources and types of bias that may invalidate the evaluation findings.

Because the goal free model is designed to release the evaluator from the constraints and biases, real or perceived, of the goals or objectives established for the program, it may be particularly suited to external evaluators who want a methodology that will enable them to judge a program's merit and worth, but

which does not force them into the thought paradigms of program developers, managers or administrators (Scriven, 1987). Further emphasizing this advantage is the idea that shifts in program goals, midway into a program, have the potential to interfere with or complicate the job of evaluating when models other than goal free are used (Scriven, 1987).

Stufflebeam (1987) says that goal free evaluations are useful in judging alternative program strategies because they would be better able than goals-based models to assess a program's worth. In this respect, the goal free model would serve in a decidedly consumerist fashion. By comparison, product evaluations done by the Consumers Union are done on products which are bought on the open market. Their aim is to judge a product's value to the consumer without being influenced by the intentions of product manufacturers, distributors or advertisers.

Disadvantages.

The foundation of the goal free model is that it operates free of the goals or objectives intended by program developers. According to Scriven (cited in Salasin, 1974), the difficulty in remaining ignorant of intended program goals increases as the size of the program and time involved in evaluating a program increases. Thus, if the evaluator's objective is to evaluate for all program effects, without being aware of those that were intended, the objective will likely be defeated if the model is

used for large programs.

For similar reasons, the model is useful only for external evaluators. Internal evaluators are likely to be too close to the program to avoid being aware of, and influenced by, the intended program goals.

Many program evaluation theorists believe that it is important for evaluators to investigate such things as interpersonal conflicts between program audience, hidden agendas and the like (Brookfield, 1986). These are sources of information which goal free evaluators must avoid because they inform the evaluation of the intended program goals. Thus, the model must necessarily impose an epistemological dilemma for those evaluators who wish to avoid the influence of intended program goals but who see merit in investigating program politics.

The methodology for the goal free model is less well defined than in the more widely used evaluation models (House, 1980). For this reason, it may not be suited for evaluators who have little training in the field of evaluation. A further and related consideration is that because the model has not been widely tested, there is a significant shortage of normative data from which to derive important operational strengths and weaknesses.

Borg and Gall (1989) note that there are likely to be constraints on the use of the goal free model because evaluators are very often employed to determine whether program goals have been met. Where an assessment of whether program goals have been achieved is prescribed as even a minimum condition for an

evaluation, and where evaluations are tendered as contracts, this model will not compete favourably with those that can meet the minimum condition.

Art Criticism

The art criticism model, alternatively known as the connoisseurship model, derives from the tradition of criticism within the field of art and literature (House, 1978). One of the more recent evaluation models, art criticism relies heavily on the insights of individuals who have expert knowledge about a precise area (Stufflebeam & Webster, 1983).

Art criticism as a model for the evaluation of educational programs was pioneered by E.W. Eisner and is substantively an elite and qualitative approach to providing the consumer, who could be anyone from layman to connoisseur, with critical judgements about the quality of a program. Stufflebeam and Webster (1983) state that the purpose of a connoisseur-based study is to "describe critically, appraise, and illuminate the particular merits of a given object" (p.35).

While art criticism is like Michael Scriven's goal free approach in that both can be described as consumerist, both are very different in terms of methodology and their concern with bias. Indeed, with regard to bias control, these two models can almost be described as polar opposites, where art criticism is the model with the most inherent bias.

Unlike the models described thus far, art criticism is essentially qualitative (Borg and Gall, 1989). This is not to suggest that these earlier described models cannot have qualitative components. It means only that art criticism evaluation is by nature qualitative in that it is "the illumination of something's qualities so that an appraisal of its value can be made" (House, 1980, p. 32). The ultimate purpose, the outcome, of such an illumination is improved standards (House, 1978).

Advantages.

One of the major advantages of the art criticism model is that it exploits an evaluator's superior, esoteric level of understanding about a specific area for the benefit of a lay audience. It is particularly beneficial for programs where the expert is highly respected within his or her field and where the audience has considerable confidence in that person's ability to provide an illumination of the nature and value of the program.

This model of evaluation also has the potential for increasing the credibility of the program type as well as the field of study with which it is associated. The impression conveyed if an evaluator who, for example, is known more for his or her work in theoretical physics is hired by a group of administrators to evaluate a reading program for elementary school children might be that the field of reading and thereby

the program is based not on any scholarly tradition but on discernable common sense. However, if the same administrators were to employ a highly regarded specialist in reading to evaluate the same program, the likelihood is greater of conveying an impression that the program, and thereby its developers, is sufficiently linked to a professional field to warrant a judgement from an expert in that field. Thus, it is a model that can appeal to the sensitivities of professionals from the field with which the program is associated, as well as inspire confidence in the general field.

House (1980) says that "properly executed criticism will increase awareness and appreciation" (p.33). By increasing the consumer's awareness and appreciation, the evaluator makes the consumer a more critical judge of other like programs. In this respect, the findings from art criticism evaluations are more apt to have greater generalizeability than evaluation studies that are more quantitative in nature.

Disadvantages.

One of the greatest weaknesses of the art criticism model is that because it relies on the expertise of the evaluator, it places the evaluation at a high risk for bias and corruption (Stufflebeam & Webster, 1983). Indeed, while the same may be said of other models, it is probably truer of art criticism that the

desired findings can be achieved by selecting the right evaluator.

Expertise is something which can be defined as a theoretical construct but its applied value is sometimes difficult to ascertain because experts in the same field can very often have widely divergent points of view. There are numerous examples from the use of expert opinions by the judicial system which demonstrate that such opinions cannot be relied upon to be consistent between individuals. Hence, any illumination about the nature and value of a program that a given expert may provide is at risk of being perceived as having value only in relation to the particular pedantic tradition in which the evaluator has been schooled.

Accreditation

The accreditation model is similar to the art criticism model in that it relies on expert or professional knowledge for a judgement about a program. Its methods, and thereby the model itself, are classed as qualitative (Patton, 1980). House (1978) says that the outcome of an evaluation using the art criticism model is improved standards whereas the outcome of an evaluation using the accreditation model is professional acceptance. And while professional acceptance also implies improved standards, a further implication is that there is an important distinction between ideal and acceptable standards and that the ideal may be

laudatory but impractical and possibly unachievable. Thus, what is professionally acceptable may be less than the ideal, but more achievable.

Worthen and Sanders (1987) describe two different types of accreditation, namely, institutional accreditation which addresses an entire institution, such as a university, and specialized or program accreditation which addresses a component or program within such an institution. They note that specialized or program accreditations are "usually more specific, rigorous, and prescriptive than those used in institutional accreditation" (p.107). For the larger purpose of this study, however, the focus of this discussion will be on program accreditation.

A characteristic feature of the accreditation model of evaluation is that it involves multiple professionals acting together as a committee or a team. Accreditations sometimes involve a preliminary internal program review in addition to on-site visits by a team of external examiners, and they generally rely on published standards (House, 1980; Worthen & Sanders, 1987). However, accreditations have also been completed by a board or council of professionals who, without the benefit of an on-site visit, arrive at an overall judgement based on forms completed and submitted to the board or council by administrators and managers of the program seeking accreditation (House, 1980).

Advantages.

One of the strongest advantages to this model is that it relies, to a large extent, on pre-established standards which a professional body, the accrediting agency, has defined as minimal requirements for the professional and ethical operation of a program of the type seeking accreditation. The accrediting agents, the evaluators, are therefore not placed in the position of having to define standards every time such an evaluation is begun.

Another advantage of accreditation is that while accreditation is not always a requirement for a program to operate, it is a symbol of achievement that is highly coveted by most programs for which it is available (Worthen and Sanders, 1987). This not only encourages program administrators and managers to maintain an ongoing concern for accreditation standards, it is also less apt than most other evaluation approaches to stimulate a defensive response from those who are perceived, or who perceive themselves, as accountable for the program.

Accreditation is also valuable in that it provides program administrators and the consumer with normative feedback. That is, it informs them about the program's performance in relation to other programs of its type, in addition to performance in relation to the standards. The potential exists, therefore, to stimulat a healthy competitiveness between programs which is

beneficial to both the consumer and the professional field with which the program is associated. Worthen and Sanders (1987) note in this regard that "most would agree that accreditation has played an important role in educational change" (p.101).

Disadvantages.

House (1980) has stated:

> Public disenchantment with professionally controlled evaluatior is reflected in the declining credibility of accrediting agencies. At one time, it was sufficient for an institution to be accredited by the proper agency to be assured of its quality - but no longer....Certainly, the public is justified in questioning some of the evaluation procedures. Visits by accrediting teams are highly variable, and their rigor depends on the members of the team" (p.238).

Although it is difficult to tell how pervasive, the perception that accreditation is an incestuous system which serves only the interests of the profession, is real and deserves consideration. Evaluations which, from the beginning, are perceived to be biased against an accurate judgement, will be susceptible to post-evaluation criticism and dismissals. Thus, accreditation, more than most models of evaluation, ought to be carefully considered for the perception of its credibility with

the consumer, the public, as well as the program audience before it is begun.

A greater disadvantage with accreditation, however, is that since accreditation is an ongoing process that is designed to judge programs which are affiliated by purpose and operation, and by membership in a regional, national or international professional standards association, it would not be suited to programs which are unique and which hold no such membership. Moreover, because it typically requires the assistance of a team of external assessors who are professionals in the field with which the program to be accredited is associated, the accreditation model may be more costly to implement than models which rely on a more limited resource of expert opinion or which are not dependent on such opinion.

Adversary

The adversary model, also known as the quasi-legal model, of evaluation is exceptional to most other models of evaluation because it begins with a specification of two polarized results that could be produced by the evaluation. That is, it begins by stating that the results indicate an excellent program, *and* that the results indicate an ineffective program. The evaluation organizers then designate two teams; an advocacy team to substantiate and argue for the positive result and an adversary team to do the same for the negative result. The teams present

their findings to a jury for a judgement (Patton, 1980). Thus the adversary model resembles a court trial where a defence and a prosecution each present and argue their evidence about a defendant. House (1980) says that the adversary model is also comparable with appointed commissions of enquiry and points out that such commissions are often not adversarial in nature. The objective for an adversarial evaluation is to make an informed judgement about a program by analysing the rigorously researched pros and cons of a program.

Compared with the other models discussed thus far, the adversarial model is the most democratic. It is a qualitative approach to investigation, and subjective in terms of research methodology.

The adversary model is classified as being both participatory and transactional. "[T]ransaction models involve people through negotiation, [and] interviewing" (House, 1978, p.10). Participatory models attempt to engage direct participation of people who have direct involvement in the program. In the adversary model, this is achieved by allowing program participants representation at a mock-trial (House, 1978).

Developed by T.R. Owens and R.L. Wolf in the early 1970s, the adversary model was designed with the intention of improving decision making in education (Owens & Wolf, 1985). However, it is one of the least used of the models discussed thus far.

Advantages.

It is argued that the adversary model is able to provide the decision maker with a higher quality and broader scope of information than most other evaluation models. The methodology, which involves at least two separate individuals or groups of people researching and building a substantiated argument for their mutually opposing positions, attempts to develop and/or harness biased groups so that the advantages and disadvantages of a program can be fully explicated (Owen and Wolf, 1985). In a court trial system, for example, it is believed that truth and justice could not be properly served if the function of defence and prosecution were provided by the same individual. The reasoning is that once an opinion is formed as to guilt or innocence, (in this case, effectiveness or ineffectiveness), the balance of effort to substantiate both for and against that opinion shifts in favour of either the for or the against. However, when the effort to substantiate is dedicated to a single pre-stated conclusion, the thoroughness of the effort to substantiate that conclusion will likely be greater than the thoroughness of the effort to substantiate multiple and opposing conclusions. Thus truth is better served, and decisions more confidently made, when the responsibility to investigate the veracity of opposing sides is divided and segregated between assigned investigators.

A further advantage of the adversary model is that it is

effective at exposing and clarifying hidden assumptions that may exist about such things as the political, social, and educational functions of the program (Patton, 1985). The opportunity to argue a position is accompanied by the right of an opposing side to question that position. Thus, through the process of the presentation of substantiating evidence and subsequent debate and questioning, the evaluating jury is positioned to developed an informed judgment of the effectiveness of the program.

Disadvantages.

Popham and Carlson (1983) cite many and significant disadvantages to the adversary model. They say first that an effective adversary evaluation is contingent on the skill and ability with which both sides present and defend their positions. Where there are inequities in such skills and abilities between individuals on either side of an issue, the stronger side will carry greater influence. Popham and Carlson (1983) give as an example the situation which can occur in court trials where affluent defendants can afford to hire experienced, top-ranked trial lawyers while more destitute defendants must settle for public defenders who are often inexperienced. The unfortunate result of this situation is that the more experienced attorney will invariably have greater success. Thus, findings from evaluations that are structured on the adversarial model risk being biased in favour of the side with the greatest research and

verbal prowess.

Another disadvantage is that the advocate-adversary conflict places programs in the position of being judged in a simple good or bad, accept or reject framework (Popham & Carlson, 1983). The problem with this is that the adversary model is suitable only for those programs for which clear, simple, dichotomous choices are required. It would not be suitable for programs for which information about changes required to affect program improvement is desired.

The adversary model is also more expensive to implement than more conventional evaluation studies. The extra costs arise from the need for two teams, advocate and adversary, and arbitration staff. It is also expensive because "there typically will be a greater demand for data because the two teams will often grasp at straws, any old straws, in an effort to bolster their case" (Popham & Carlson, 1983, p.211).

Transaction

House (1978) describes the transaction model, also known as the case study model, as the most democratic of all models represented in his taxonomy of major evaluation models. An inherently qualitative approach, it solicits the opinion of a broad cross-section of people who have involvement with the program being evaluated and attempts to provide findings which reflect those opinions. It does not ignore individuals or groups,

as is the case with some models, simply because their opinion is not deemed important in relation to the process of measuring the achievement of predefined program objectives.

Methodologically, transactional evaluations usually involve interviews with program audiences and on-site observations (House, 1980). The transactional "approach concentrates on the educational processes themselves: the classroom, the school, the program. It uses various informal methods of investigation and has been drawn increasingly to the case study as the major methodology" (House, 1978, p.5).

House (cited in Patton, 1980) says that the transactional model reflects a subjectivist epistemology and usually follows a naturalistic inquiry methodology. Other models which House classifies as having a subjectivist epistemology are art criticism, accreditation, and adversary.

Evaluation models deriving from a subjectivist epistemology define the value of a program "based on personal judgement" which means that "[t]here is a good deal of subjectivity in how the key actors conduct their investigations" (House, 1978, p.5). House (1978) says further that "[t]he subjectivists are less interested in arriving at a proposition that is "true" (in the generalizeable sense) than in relating the evaluation to the particular experience of the audience" (p.8).

The naturalistic inquiry methodology engages program participants and stakeholders as if they are collaborators in a process which culminates in a judgment about the program

(Williams, 1986). However, as Fetterman (1986) points out, naturalistic inquiry is a generic term which describes various types of qualitative studies. Ethnographic evaluation, for example, which attempts to "understand and describe people's perceptions of reality", is one of the more prominent kinds of inquiry which use the naturalistic methodology (Fetterman, 1986, p.23). And ethnographic evaluation is quintessentially subjectivist in epistemology.

> The transition [sic] model is based on the same assumptions that undergird qualitative research: the importance of understanding people and programs in context; a commitment to study naturally occurring phenomena without introducing external controls or manipulation; and the assumption that understanding emerges most meaningfully from an inductive analysis of open-ended, detailed, descriptive, and quotive data gathered through direct contact with the program and its participants (Patton, 1980, p.55).

Among the principal proponents of the transactional model is Robert Stake whose "responsive evaluation" approach has become the leading example (Patton, 1980; House, 1978; House, 1980). Evaluators using the responsive approach usually negotiate "with the client as to what is to be done...[and respond] to what different audiences want to know" (House, 1980, p.40).

Stake (1983) says that while the responsive evaluation

approach is original, it is based on an old ideology; namely, it is based on the idea that evaluation involves observing and reacting, things people would be naturally inclined to do. It is an approach which does not rely on "preordinate" plans which itemize and describe program goals, which use objective tests, and which use standards acceptable to program personnel as an evaluation guide (Stake, 1983). Rather, it is "an emergent form of evaluation that takes as its organizer the *concerns and issues of stakeholding audiences*" (Guba & Lincoln, 1981, p.23).

Stake (1977) says:

> An evaluation is *responsive evaluation* (1) if it orients more directly to programme activities than to programme intents, (2) if it responds to audience requirements for information and (3) if the different value-perspectives present are referred to in reporting the success and failure of the programme (p.163).

Stake believes that the evaluation organizers, the standards and criteria against which the evaluator should judge the program, should emerge from the concerns and issues of the stakeholding audiences, and that these concerns and issues should be gathered from conversations with persons associated with the program (Stake, 1975). He believes that evaluations can serve many purposes, but that the purpose for any given evaluation

should be defined by the purposes and information needs of the different audiences (Stake, 1977; Guba & Lincoln, 1981).

The rationale for defining the purpose of an evaluation in terms of the purposes and information needs of the different stakeholding audiences is that it increases the usefulness of the findings to the people associated with the program. The trade-off to using this approach, however, is that it tends to "sacrifice some precision in measurement" (Stake, 1983, p.292).

Stake developed a simple, heuristic diagram (see Figure 2) to help describe the 12 recurring, prominent events that comprise a responsive evaluation. Although the twelve events in the diagram are laid out in the form of a clock, the events themselves are not to be read in an exclusively clockwise fashion. Rather, it ought to read in a hypertext fashion, which could mean clockwise, counter-clockwise or cross-clockwise.

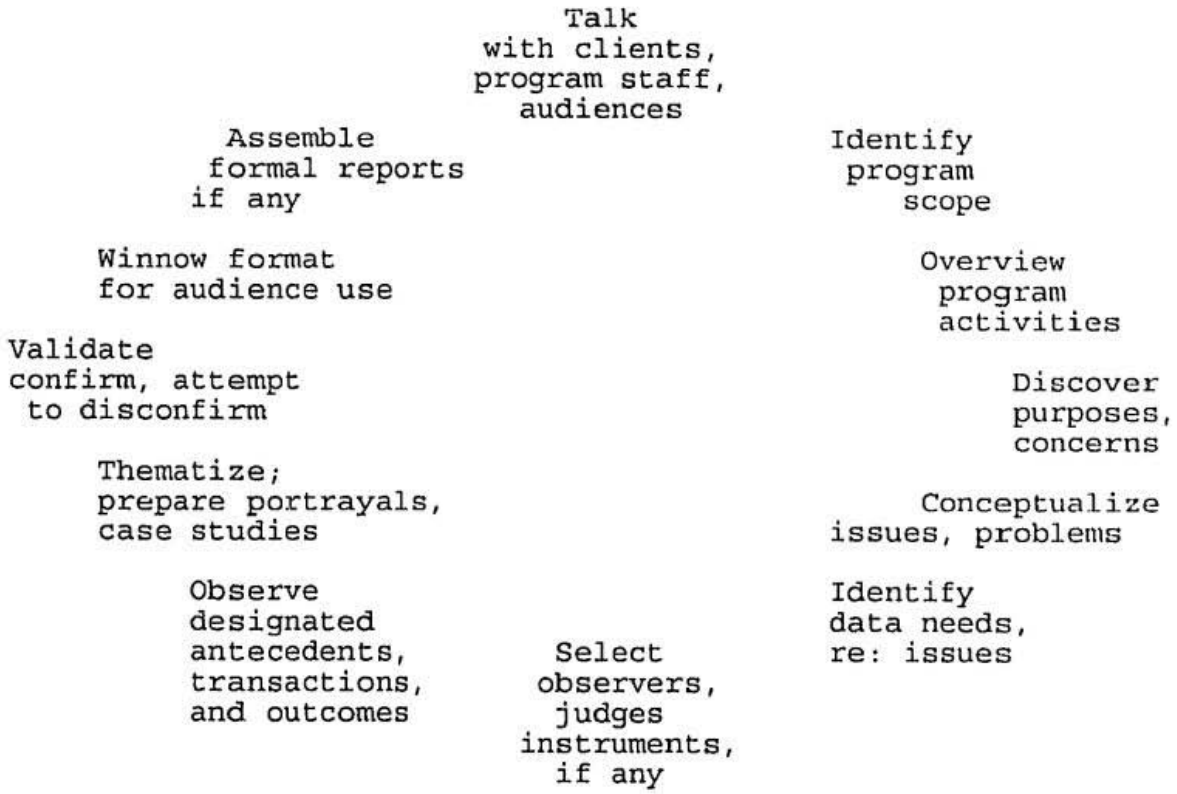


Figure 2. Prominent events in a responsive evaluation (Stake, 1983, p.298).

The following discussion of advantages and disadvantages of the transactional model will be in relation to Robert Stake's responsive evaluation approach, since it is the most widely known and used example.

Advantages.

Guba and Lincoln (1981) state that unlike earlier models for program evaluation which took the "singular value perspective"

that all value systems converge to be consensual, responsive evaluation took a "pluralistic view", allowing the possibility of conflicts between value systems that do not necessarily converge (p. 29). Thus, responsive evaluation is advantageous for evaluation sponsors who are concerned with informing program audiences about a program's value from the perspective of the "multiple realities" generated by the concerns and issues of each audience or audience group (Guba & Lincoln, 1986).

The responsive approach is valuable for its flexibility. The method by which the evaluator communicates the findings is not pre-defined. It does not have to conform to the classic research-type report associated with most of what Stake calls "preordinate" evaluation models; namely, those models which are typified by the use of goal statements and the use of objective tests. Instead, the responsive approach allows the evaluator to choose "media accessible to his audiences to increase the likelihood and fidelity of communication. He might prepare a final written report, he might not - depending on what he and his clients have agreed on" (Stake, 1975, p. 14).

House (1978) calls the transaction model, of which Robert Stake's responsive approach is representative, as being highly democratic. MacDonald (cited in House, 1980) says that,

> Democratic evaluation is an information service to the community about the characteristics of an educational program. It recognizes value-pluralism and seeks to

represent a range of interests in its issue formulation. The basic value is an informed citizenry, and its evaluator acts as broker in exchanges of information between differing groups. His techniques of data-gathering and presentation must be accessible to non-specialist audiences. His main activity is the collection of definitions of, and reactions to, the program." (p.41).

Thus, where evaluation sponsors desire an evaluation that serves and speaks to the community at large, evaluation approaches such as the responsive approach are highly favourable. The responsive evaluation approach, like other transactional approaches, recognizes that some programs more than others hold great interest to many individuals and groups within the community, and that any effort to measure such a program should focus on considering and responding to the concerns and issues of those individuals and groups.

The responsive evaluation approach is also advantageous for evaluating programs for which the client or sponsor's needs have not been articulated at the beginning of the program (Kirkup, 1986). Because the responsive evaluation approach considers the client's needs as no more or less important than the needs of other program audiences, and because the approach is designed to allow the needs of all program audiences to emerge during the evaluation process and still provide useful results, it is beneficial for use in evaluating programs for which an evaluation

is deemed necessary, but which lack defined standards and criteria for making an evaluative judgement.

The responsive evaluation approach is valuable for use with community-based programs where it is seen to be important to make all program audiences feel that they are a legitimate and important part of the evaluation process. The methodology for the responsive approach involves observations and interviews. It requires a high level of interaction between the evaluator and the audiences involved in the program, and thereby gives all audiences a sense of control and ownership of the evaluation.

Guba & Lincoln (1981) state that they believe that "responsive evaluation as proposed by Stake and elaborated by others offers the most meaningful and useful approach to performing evaluations" (p.33).

Disadvantages.

One of the potential weakness to the responsive approach is that it is less precise in measurement than preordinate and less qualitative approaches. The argument to defend this lack of precision is that it will "hopefully...increase the usefulness of the findings to person (sic) in and around the program" (Stake, 1983, p. 292).

A further potential weakness is that because the responsive approach is based on the philosophy of responding to needs of the various audiences, the needs of some audiences or audience groups

may dominate by virtue of the fact they are better able to assert and articulate their needs than others (Logsdon, Taylor & Blum, 1988). Where one audience, such as program participants, for example, is dispersed over a wide geographical area and another, such as program developers, is localized, the ease with which the evaluator is able to interview both groups is unequal. Thus, the notions of credibility and bias can come into question.

Another potential disadvantage of the responsive approach is that its methodology is heuristic, which means that evaluators do not have a clear, procedural or step-by-step path to follow. This may present difficulties, particularly for novice evaluators, because the competing needs of the multiple audiences can place extreme demands on the organizational, information management, and negotiation skills of the evalautor (Sadler, 1981).

Chapter 3

Methodology

Evaluation Model Selection

The Distance Education Course for Literacy Providers was a response to the altruism and energy of Newfoundlanders concerned with the social and economic consequences of their complacency in the face of an intolerably high number of fellow adult Newfoundlanders blighted by illiteracy. For this, and numerous historical reasons, it had sociocultural significance for Newfoundland and Labrador.

Until the turn of the twentieth century, Newfoundland's economy rested almost exclusively on a single resource: the commercial fishery. It was characterized by a truck system controlled by merchants who established the value, to be given as credit toward merchandise in their stores, for each fisherman's catch. The effective result was an oligarchy where the powerful few held the overwhelming majority in economic and political submission.

Newfoundland's population is highly homogeneous. The people descend from illiterate English and Irish peasant workers, most of whom emigrated to the Island in the period following the Seven Years' War (1756-63) and the American Revolutionary War (1775-83) (Handcock, 1989). Moreover, the pattern of settlement was one which did not lend well to the development of infrastructures for transportation, communication, health care or education. Of the

nearly 1400 communities that chained along the Island's 10,000 kilometres of isolated coastline in 1901, 59 percent had a population of less than 100 and 39 percent had a population of less than 25 (McCann, 1988). In short, Newfoundland became a time capsule that preserved and perpetuated a subsistence lifestyle, a predominately oral communication tradition, and a politically subordinated peasant culture.

Newfoundland's economic and demographic profile began to improve slowly after Confederation with Canada in 1949, but the embedded social and cultural traditions that had developed over the 300 years prior were very slow to change and adapt to ways of the sophisticated western industrial world. Two of the most significant manifestations of this resistance to change can be found in the facts that in the period between 1961 and 1989 the number of Newfoundland fishermen, as a percentage of the total population, increased from four to five percent (Historical Statistics, 1990), and that the Island's rate of functional illiteracy, measured by survey in 1987 as being 44 percent, was the highest of all the Canadian provinces (Southam Newspaper Group, 1987).

Elsey (1993) says:

> The notion of social purpose draws its inspiration from the values of an active and participatory democratic citizenship and the need for informed critical discussion as the basis of a vibrant political culture using knowledge as social and

personal power to deal with the rapidity of change in modern society. These ideas about adult education for social purpose have waned as a central political value in economic rationalist and complacent democracies, regretfully, but trace elements are still alive in some corners of adult education, most notably at the level of neighbourhood and community action concerned to further a cause through exerting political pressure" (p.10).

The Distance Education for Literacy Providers (DELP) Course was a response to, and an opportunity to further engender, social purpose. It was a pilot project intended to promote active and participatory democratic citizenship and to thereby empower those involved at the community and neighbourhood level with the knowledge and skills to attack the problem of adult illiteracy. In this respect, the community action that the DELP Course was designed to assist was reminiscent of the self-education societies that emerged during the 1770s and 1780s in England. Unlike the self-education societies, however, which were officially banned, the community-based action on the part of literacy providers was recognized and supported by the DELP Course.

The decision on the part of the evaluators to choose Robert Stake's responsive evaluation approach was supportive of the notions of social purpose and participatory democratic citizenship that were symbolized by the Distance Education for

Literacy Providers Course. It would offer program participants, the neighbourhood and community literacy providers for whom it was designed, as much say in determining the issues and concerns on which the evaluation would focus as it would to any other audience group. It would therefore provide the evaluation sponsors, as well the other identified audiences and audience groups, with results that would not only demonstrate wide consultation, but would also offer feedback for improving the DELP Course to make it suit the practical realities of life, work and economics for literacy providers in rural communities and urban neighbourhoods in Newfoundland and Labrador.

The methodology for the responsive approach, which involves interviews, and observations, would bring the evaluators close to all audiences and thereby communicate that the evaluation was intended as a participatory exercise, without elitist intentions. It would also communicate and demonstrate that subjectivity in evaluation is as epistemologically valid as objectivity, and that a subjective methodology would produce more obviously direct links between their concerns and issues and the evaluation outcomes.

In short, the responsive approach to evaluation approach would suit, better that any of the other models discussed in the previous chapter, the historical, cultural and emancipative spirit in which the DELP Course was conceived and the social context in which it would be delivered.

Methodological Procedures

The evaluation of the DELP Course was undertaken using a modified version of Robert Stake's Responsive Model. The original Stake model has twelve phases diagrammatically arranged in the shape of a clock. The modified version (see Figure 3) contains eight phases which, like the phases in Stake's original diagram, are iterative in that they can be read and followed in a clockwise, counter clockwise and cross clockwise fashion.

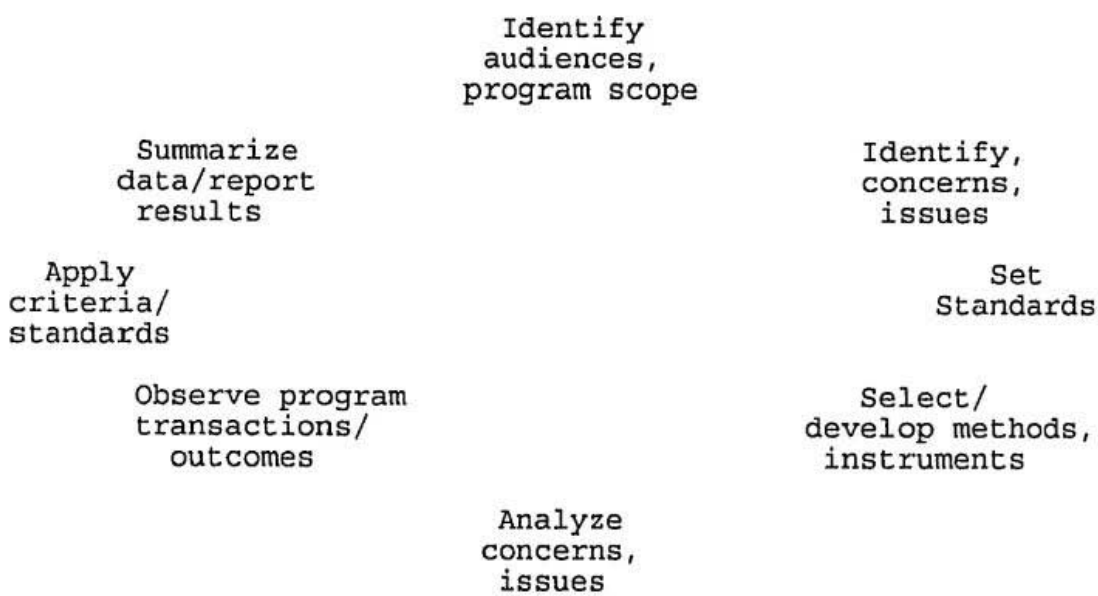


Figure 3 An adaptation of Stake's prominent events in Responsive Evaluation (Stake, 1976) to the Distance Education Course for Literacy Providers Evaluation (see Lertpradist, 1990, p.99)

Audience Identification and Consultation

Through a series of preliminary interviews beginning with the more prominent audiences, all program audiences were identified. Semi-structured interviews and short telephone administered questionnaires were used to gather the concerns and issues of all identified program audiences.

Three separate audience groups were identified by the evaluators. Without any intent to prioritize, the first audience group identified was the Curriculum Committee. While the committee was interviewed as a group, and one individually, for the purpose of gathering their concerns and issues, the committee itself was transient. Time constraints and employment commitments on the part of committee members and program budget constraints meant the committee had to conclude its activities quickly. Nevertheless, all members were later consulted for their approval of the established program standards and criteria.

Course participants, subdivided according to previous training, formed the second audience group. Thirty in total, they were all contacted by telephone for the purpose of gathering relevant biographical data as well as concerns, issues and expectations they had about the course. The purpose for subdividing this group was to supply the evaluators with data that would enable them to judge the effect of the course on participants with differing amounts of experience and training.

The third group, the course/program sponsors (federal and

provincial governments), was represented by the director of the provincial government's Literacy Policy Office. This person was contacted by telephone for the purpose of determining concerns and issues the sponsors had about program design and implementation and to gain additional background information.

Concerns and Issues Analysis

Guba and Lincoln (1981) define a concern as "any matter of interest to one or more parties", while an issue is defined as "any statement or focus about which there are different points of view, or any point of contention" (p.92). There were no issues identified among audience groups. The following is the list of main concerns identified:

- Gains in tutor knowledge and self-esteem
- Knowledge gain and positive attitudinal changes for both "new entrants" (i.e., those with little experience and training) and "literacy personnel" (i.e., those with extensive experience and training);
- Versatility of program and materials in terms of delivery mode;
- Relevance of program and materials for intended audience;
- Efficacy of the delivery mode;
- The combination of media used;
- Cost implications of the teleconference approach.

Document Analysis

All documents and program materials associated with the development and implementation of the course were obtained and analyzed for their quality and relationship to the components of the course, and for their ability to assist the evaluators in the creation of evaluation standards and criteria. The documents included a project proposal for a community-based adult literacy resource kit; the proposal for the design of the DELP Course; a program and activity guide; an unedited collection of participant and resource person biographies; a course manual; a recently written and published autobiography of a Newfoundland woman whose adult life had been transformed by newly acquired literacy skills; and three two-hour program video tapes.

Evaluation Standards

The evaluators used the concerns of the audience groups, and program goals and objectives which were gleaned from course documents, to develop standards and criteria for the evaluation. Once created, the evaluation standards were communicated to the audiences for approval. They were given the opportunity to object or make modifications to the developed standards. All standards were accepted as valid and acceptable measures for the evaluators to employ in the process of making a judgement about the program. It was noted that not all standards could be measured during the

implementation of the course. Several standards would require follow-up with DELP Course participants. The following are the standards and criteria which were developed by evaluators:

Standard 1. The curriculum for this program should satisfy participant needs.

Criteria 1. The program should provide participants with increased knowledge of principles and methods of literacy provision in accordance with their stated needs.

Criteria 2. The curriculum should meet the expectations of the various audiences.

Criteria 3. The curriculum has the potential to achieve stated program goals and objectives.

Criteria 4. The curriculum has the variety and scope to satisfy the needs of both groups of participants, i.e. those with and without previous training.

Criteria 5. Program participants find the curriculum sufficiently informative to want to complete the program.

Standard 2. Program objectives should be clearly stated in writing.

Criteria 1. Program objectives should be clearly stated in the course text or otherwise be clearly and explicitly disseminated to program participants.

Criteria 2. Program objectives should meet the expectations of the various audiences.

Standard 3. Participants who complete this program should be capable of applying, to their tutoring practice, the knowledge and skills which the curriculum intends to provide.

Criteria 1. Program participants should be observed assimilating or (at a minimum) attest to the use of knowledge and skills learned in the program.

Criteria 2. Program participants should be capable of stating what knowledge and skills they acquired from the program.

Standard 4. Knowledge and skills which the participants acquire through this program will benefit both the program participants and the low-literate learners.

Criteria 1. Knowledge and skills which the participants acquire through this program, when applied in the learning environment, affect positive change on the facilitation of learning.

Standard 5. Program goals and objectives, given program constraints, should be both achievable and feasible.

Criteria 1. All goals and objectives are met at the conclusion of the program.

Standard 6. Program provides opportunity for sufficient participant discussion and the sharing of ideas.

Criteria 1. Program has regular time scheduled in each teleconference session for participant interaction.

Criteria 2. Program includes activities designed to encourage and facilitate participant discussion.

Criteria 3. Teleconference moderators/leaders should orchestrate activities which promote participant questions and discussion.

Criteria 4. The opportunity for discussion provided in the teleconference sessions meets the expectations of program participants.

Standard 7. Program participants possess a post-program improvement in their self-confidence of their abilities as literacy providers.

Criteria 1. Program participants express, verbally, a perceived improvement in their self-confidence of their abilities as literacy providers.

Standard 8. Curriculum materials and teleconference sessions should be presented in a manner consistent with the level of prior knowledge and training of participants.

Criteria 1. Program participants, at the conclusion of the program, feel that both the curriculum content and the method of presentation was neither too advanced nor too elementary and with sufficient respect for dignity.

Standard 9. The combination of media used for the delivery of this program should be suitable for both the curriculum content and the participants.

Criteria 1. Text based materials should be clearly modularized to correspond with the weekly teleconference sessions.

Criteria 2. Program participants should find the text based materials easy to read, attractive in format and interesting.

Criteria 3. Program participants should find the videotaped course materials to be interesting, easy to understand and relevant to teleconference discussion and text based materials.

Criteria 4. Program participants should find teleconference sites to be easily accessible.

Criteria 5. Program participants should find the frequency of teleconference sessions to be suitable for effective learning.

Criteria 6. Program participants should find the duration of each teleconference to be suitable for effective learning.

Criteria 7. The delivery system should meet the expectations of program participants.

Standard 10. The course, designed for the teleconference delivery mode, with print and videotape support materials, is suitable for delivery through other modes.

Criteria 1. The teleconference course is instructor independent and/or suitable for audiotaping and/or packaging.

Criteria 2. Program participants, upon completion of the program, are able to deliver components of the program through classroom-based instruction.

Criteria 3. All support materials for the program are complete, clearly linked to program modules and are suitable for supporting delivery of separate program components through classroom-based instruction.

Observation Data Collection

The DELP Course was offered by teleconference one night per week over a ten week period to 15 locations on the Island and in Labrador. Each weekly meeting was attended by one of the

evaluators. Observations for nine of the weekly meetings were made in classrooms at the source teleconference centre in St. John's. For the penultimate meeting, the evaluators travelled to three of the remote teleconference locations (i.e., Gander, Corner Brook & Stephenville). Observations made from the remote locations enabled the evaluators to assess the conditions at these sites. Being at the remote sites also enabled the evaluators to complete a total of eight pre-arranged, in-depth interviews with course participants located at these sites.

The evaluators developed a Record of Observations form to assist them in recording observational data that related to the evaluation standards and criteria. An attendance sheet was also developed to track course participant attendance; as part of the weekly teleconference sign-on procedure, participants at each site were asked to give their name and location.

Evaluation Instruments

Five different instruments were developed to collect data from the audiences. The instrument used to gather preliminary information and the concerns and issues of the Curriculum Committee and the literacy program sponsors was a semi-structured questionnaire containing five open-ended questions. The instrument used to gather preliminary information and the concerns and issues of the literacy providers contained 17 open-ended questions. Both of these instruments were administered over

the phone.

The instrument used to gather data about the literacy providers' experience and feelings about the program within a month of its completion consisted of 51 forced-choice questions and five open-ended questions. This instrument took about 15 minutes to complete, and while most of the interviews were completed by phone, a few were done in person. The instrument used for the in-depth interviews with eight of the literacy providers consisted of nine open-ended questions. These in-depth interviews were done face-to-face and were completed by the evaluators during a visit to the three remote teleconference locations cited above. Together with observation data gathered by the evaluators, the post-program questionnaire and the in-depth interviews provided data with which to measure eight of the ten evaluation standards (i.e., 1 & 2, & 5 through 10).

The final instrument was used six months after the DELP Course had finished and helped the evaluators gather data from the Course participants to measure the remaining evaluation standards (i.e., 3 and 4). It consisted of seven closed-response items, four open-response items, and nine statement items to which respondents were asked to reply using a four point Likert scale. This instrument was also administered by phone (see Appendix B for all instruments).

Data Analysis

The data for the evaluation were collected in four stages, each stage serving a different function. Data gathered during the first stage were entirely qualitative and served the function of enabling the evaluators to establish evaluation standards and criteria. Program audiences were identified and interviewed and their responses classified so that all concerns and issues could be converted into standards. Once standards had been derived, the evaluators generated criteria with which to measure the degree to which they had been achieved. The criteria also formed the basis for the development of further evaluation instruments and observation guides.

The remaining three stages of data gathering were characterized by (1) observations made during the weekly teleconference meetings; (2) the administration of a post-program questionnaire to all course participants, and an in-depth interview with eight participants; and (3) the administration to all course participants of a nine item questionnaire requiring a Likert scale response.

Data from the final three stages of data collection were a combination of forced choice and open-ended responses. With the exception of data from the in-depth interviews, all data were classified in relation to the evaluation standards and criteria. Forced choice responses were organized in table format while open-ended responses were quoted in part or in entirety and

listed under the heading of questionnaire items which prompted them.

Responses from the in-depth interviews were abstracted and listed in an appendix of the Interim Report. Respondents were identifiable only by code number.

CHAPTER 4

The Results of the Study

Introduction

This chapter is organized around the standards developed by the evaluators, and approved by the audiences. Data from analysis of documents, site observations, interviews, and questionnaires are summarized briefly. This is followed by conclusions drawn in relation to the standards and criteria.

Because the results of the study were disseminated in two reports, they will be correspondingly organized in this chapter. The results under the Interim Report heading will relate to standards 1, 2, and 5 through 10. The results under the Final Report heading will relate to standards 3 and 4.

Interim Report

Standard 1. The curriculum of the literacy providers' program satisfies participants' needs.

Evaluators judged that this standard was met. Participants indicated, in their questionnaire responses, that the content was of value to them, and that for the most part their information needs were met. Furthermore, the curriculum was perceived as beneficial by both groups of participants - those with prior training and those with no prior training. Both groups, based on

questionnaire responses, felt that they gained insights into literacy provision (see Tables 1 & 2).

Table 1.

Level of satisfaction with curriculum

Category of response	n=28	%
Program what I needed.	26	93%
Happy with coverage of topics.	25	89%
Wanted additional time on certain topics.	18	64%

Table 2.

Level of satisfaction with program by training

Program found to be helpful	n=28	%
With previous training (was helpful)	25	100%
Without previous training (was helpful)	3	100%

In analysing course documents (videotapes, course manual, and teleconference guide) evaluators noted that the majority of topics of specific interest, as listed by participants in their initial interviews with evaluators, were incorporated in the

course. Of 20 specific topic areas which participants would have liked to see included in the course, evaluators deemed that 13 were included. In the indepth interviews conducted with 8 participants at the end of the course, all indicated that they were very satisfied with the curriculum of the literacy provision course.

Of 30 participants who started the literacy program, only one failed to complete the course, and she indicated that she had to withdraw because of unforeseen employment responsibilities. Attendance records indicate that participants were in regular attendance, and that they managed to attend most sessions, despite winter weather conditions and, for some participants, the need to drive long distances (see Table 3).

Table 3.

Attendance

Frequency of attendance	n=29	%
Attended all teleconferences	8	27.6%
Missed one teleconference	12	41.4%
Missed two teleconferences	6	20.7%
Missed three teleconferences	3	10.3%

Conclusion

Evaluators concluded that the curriculum met the needs of both groups of program participants - those with prior literacy training and those with no prior literacy training. They concluded that no improvements are needed in the curriculum, in terms of the addition or deletion of content.

Standard 2. The program objectives should be clearly stated in writing.

In analysing program documents evaluators examined the specific goal statements and objectives as listed in the teleconference guide (agenda and handouts). They determined that the goals and objectives were clearly stated. However, the specific objectives listed for each teleconference, while serving as detailed directional statements for the participants, were not substantive enough to serve as program objectives (i.e. statements indicating learner outcomes). Evaluators noted, however, that the goal statements for each of the teleconferences actually indicated learner outcomes, hence they treated the goal statements as specific program objectives.

Program objectives met the expectations of various audiences. The curriculum committee ratified evaluators' standards, which incorporated the objectives. Participants indicated that they had read the objectives, that they considered them to be worth achieving, and that they had attained some or most of the objectives (see Table 4). Evaluators judged that this

standard had been achieved.

Conclusion

While the specific objective statements for each teleconference lacked the substantiveness to serve as program objectives, they served well as weekly directional statements for the participants. Goal statements were clearly stated and served the purpose of program objectives - that of indicating learner outcomes.

Standard 5. The program goals and objectives, given program constraints, are both achievable and feasible.

Evaluators judged that this standard was met. Participants indicated that they had attained objectives specific to their training needs (see Table 4). In addition, indepth interviews with selected program participants indicated that they were pleased with their learning outcomes.

Table 4.

Program objectives

Category of Response	n=28	%
Read program objectives included in materials	24	85.7%
Perceived objectives as worth achieving	26	92.9%
Felt I achieved some of the program objectives	14	50%
Felt I achieved most of the program objectives	14	50%

In analysing course documents, evaluators determined that the objectives were feasible, given the constraints of time (a ten week course), and of voluntary participation. Furthermore analysis of data from participants indicated that a significant number of objectives were achieved, and that participants were very pleased with their levels of achievement.

Conclusion

Evaluators concluded that the program goals and objectives were achievable and feasible within the ten week, voluntary parameters of the course.

Standard 6. The program provides opportunity for sufficient participant discussion and the sharing of ideas.

Evaluators judged that there was more than adequate opportunity for participant discussion and sharing of ideas. Observational records kept by evaluators indicated that the two hour teleconference sessions included regular opportunities for participant discussion and sharing of ideas. Even within presentations made by guest presenters there was no longer than a fifteen minute delivery or time span between which participants were invited to make comments or to ask questions.

The majority of participants indicated that there was plenty of opportunity for participant interaction (see Table 5). Additional comments on the questionnaire indicated, however, that while some interaction was preferred, there was too much discussion and too much repetition in the discussions. Similarly the indepth interviews indicated that participants, for the most part, felt that there was too much undirected discussion/interaction. Seven of the eight participants interviewed commented on participation/interaction, with five of the seven expressing negative opinions about the amount of interaction and the lack of organization or focus to the interaction.

Table 5.

Participant interaction at teleconference sessions.

Response category	n=28	%
Believe it is important to learn from peers	26	92.6%
Felt discussion & participation were encouraged	26	92.6%
Enjoyed opportunity for discussion	26	92.6%
Not enough opportunity for discussion	3	10.7%
Sufficient opportunity for discussion	21	75%
Too much opportunity for discussion	3	10.7%
Beneficial to hear other participants talk	24	85.7%

Evaluators, in their observation of all of the teleconference sessions, noted the same reaction as that expressed by the majority of participants. Many times during teleconferences when participants from other sites were sharing ideas and discussing particular concerns, the participants at the sites being observed would "tune out" by talking among themselves and generally ignoring the audio feed until the guest presenter or coordinator would move on to the next part of the presentation.

Conclusion

Evaluators consulted participants at the beginning of the program regarding their concerns and issues, and approximately one quarter of them indicated that they wanted the opportunity, within the teleconferences, to share and learn from each other. It seems that the sharing/discussion experiences within the sessions was a disappointment to the majority of participants, however. Evaluators concluded that the problem with participant interaction lay mainly with its lack of focus. Through activities completed by participants prior to teleconferences, coordinators hoped to focus or guide interaction, but that failed to occur. Frequently, participant interaction was in response to general verbal prompts (i.e., does anyone have any comments?). Thus while there was plenty of opportunity for interaction, the quality of discussion/sharing left the impression, for many participants, that time was being wasted and that sessions were dragging. Such participation and interaction needs to be better directed and orchestrated if the desired effect is to be achieved.

Standard 7. Program participants possess a post-program improvement in their self-confidence of their abilities as literacy providers.

Evaluators judged that this standard was met. Approximately half of the participants, when interviewed at the beginning of

the program, indicated that they would like to receive reassurance and/or confirmation regarding their literacy tutoring strategies. Because they worked one-on-one with literacy clients and had little opportunity to share tutoring tips/practices with other literacy tutors, they were unsure of their strategies and practices.

Participants were asked, through the rating scale completed at the end of the program, to provide feedback on their self-confidence levels as a result of the program. All participants rated the items dealing with self-confidence positively, averaging 3.3 on the four point scale (see Table 6. Indepth interviews also indicated that participants felt better about their ability to deliver literacy tutoring.

Table 6.

Participant rating of self-confidence

Category of Response	Rating (1 - 4)
I think I will be a better tutor.	3.3
I will not have so many doubts about what I do as a tutor.	3.3
I will feel more able to help my clients.	3.3
I have more confidence as a literacy provider.	3.4

Conclusion

Evaluators concluded that the program provided participants with confirmation regarding the efficacy of their tutoring practices, and participants felt that the program was beneficial in enhancing their self-confidence.

Standard 8. Curriculum materials and teleconference materials should be presented in a manner consistent with the level of prior knowledge and training of participants.

Evaluators judged that the course materials and teleconference sessions were suitable for all participants, despite the range of training backgrounds within the pilot group. All participants indicated, too, that the program met their specific training needs. Obviously those with significant amounts of prior training would have benefitted in a different manner than those with no prior training, but responses to questionnaire items indicated that all found the program beneficial (see Table 7).

Table 7.

Perceived benefit derived from program.

Response category	n=28	%
Content was too basic for my needs	2	7.1%
Content was sufficient for my needs	23	82.1%
Content was too advanced for my needs	1	3.6%

Conclusion

Evaluators concluded that the program materials and teleconference sessions were suitable for use with a broad spectrum of literacy providers, from those with no previous training to those with sufficient training to deliver literacy training programs.

Standard 9. The combination of media used for the delivery of the program should be suitable for both the curriculum content and the participants.

Evaluators judged that this standard was met. Evaluators analyzed all program materials and found that they were (a) comprehensive and complete, (b) well-organized, (c) relevant, (d) cross-referenced, and (e) professionally executed and presented.

In addition five participants, when asked to name the aspect of the course that they found most beneficial, indicated that the print materials were the most beneficial aspect. Indepth interviews provided evaluators with confirmation of this data. All participants commented on the utility of the course materials, specifically mentioning the value of the course manual as a future resource in their literacy provision.

There were minor criticisms of one aspect of the media used in course delivery - these were aimed at the teleconference delivery system. While most participants appreciated being able to avail of the training from their local areas, approximately half of those responding to the participant questionnaire noted occasional problems with the system, and five participants noted that such problems made the course less enjoyable (see Table 8).

Table 8.

Suitability of the teleconference system (TCS)

Response category	n=28	%
TCS was easily accessible	26	92.9%
TCS easy to use	25	89.3%
Had occasional trouble with TCS	14	50%
My problems with TCS made course less enjoyable	5	17.6%
Teleconference length (2 hours) was appropriate	25	89.3%
Teleconferences should have been less frequent	1	3.6%
Frequency of teleconferences was appropriate	22	78.6%
Teleconferences should have been more frequent	2	7.1%
Effective combination of TCS, videotapes & print	28	100%

Evaluators, in their observations at sites outside St. John's, noted the frustrations of participants when teleconference equipment did not function properly. In two of the three sites observed during the last teleconference session participants had technical difficulties, and even when they were solved readily, as was the case at one site, participants expressed dislike for the system.

Conclusion

Evaluators concluded that the combination of media was effective

and suitable for both delivery of course content and for the participants. However the teleconference system proved to be a problem for some of the participants. Evaluators realized that technical problems were beyond the scope of program coordinators to rectify, but some of the frustration might have been lessened if the teleconference guide provided clear instructions for the system, accompanied by diagrams and trouble-shooting tips.

Standard 10. The program, designed for the teleconference delivery mode with print and video support materials, is suitable for delivery through other modes.

Evaluators judged that the program was suitable for delivery through other instructional modes, such as live or on-site delivery and as a packaged course with a local coordinator. In analysing all course documents and materials evaluators found that the course manual, the teleconference guide, and the videotapes were well documented and cross-referenced. With little development the weekly teleconference presentations could be presented in audiotape or videotape formats, and in fact these presentations might even be improved through judicious editing.

Participants felt that the teleconference sessions were the least significant component of the course, in that much of the information presented by guests was already presented in the videotapes and in the course manual. As indicated in Table 9,

participants' indication of the most beneficial aspects of the course included no aspects which required the teleconference system. In addition it should be noted that the teleconference system demands a high level of skill from presenters, and the use of an alternative delivery system might not be as demanding.

Table 9.

Course aspects cited as being most beneficial

Course aspect	n=28	%
Opportunity to interact with other literacy providers	16	57%
Adult learning theory	6	21%
Guest speakers	8	28%
Sessions on spelling	6	21%
Sessions on learning disabilities	4	14%
Stages of literacy	3	11%
New sources for learning resources	7	25%
Course print materials	5	18%
Course videotapes	2	7%
Sessions on writing	4	14%
Practical knowledge for tutoring	9	32%
Sessions on reading	1	4%

Conclusion

Evaluators concluded that the program was suitable for delivery through other modes. Little further development would be required to convert the teleconference aspect of the course to audiotape, videotape, or print, or any combination of these media. And while guest speakers and participant interaction were deemed important aspects of the course, both aspects could be accommodated through other delivery modes.

Final Report

Standard 3. Participants who complete this program should be capable of applying to their tutoring practice the knowledge and skills which the curriculum intends to provide.

Evaluators concluded that participants were able to transfer knowledge and skills acquired in the program to their tutoring practice. Evaluators noted during the course that since many of the program participants were volunteers, there might be some difficulty in obtaining a measure of the achievement of this standard as volunteer work can tend to be erratic. The follow-up questionnaire, however, showed that 61 percent (17 of 28 respondents) of the program participants did work as literacy providers in the six months that followed the completion of the

Course as indicated by Table 10. Indeed, as Table 11 indicates, the average number of individuals tutored since the completion of the Course was 3.6. Moreover, of the 17 participants who were active in literacy provision, nearly all indicated that they were able to transfer knowledge and skills acquired in the Course to their practice (see Table 12). Further support for this conclusion came from responses to open-ended questions posed to participants in the follow-up questionnaire; all participants could provide specific examples of methods and resources that they had integrated, and practices such as the use of alternative instructional methods, the language experience approach, personal dictionaries, story creation techniques, and the CLOZE procedure were frequently cited (see Descriptive Data, p. 136).

Table 10

Course participants' tutoring history

Participants	n=28	%
Pre-course literacy tutor experience	26	93%
Worked as a literacy tutor since the course	17	61%
Currently working as a literacy tutor	15	54%

Table 11

Level of tutoring activity

Participants	n=17
Average number of individuals tutored since the completion of course	3.6

Table 12

Effects of course on tutoring activity

Participants	n=17	%
Feels course had positive affect on tutoring ability	17	100%
Has put ideas or methods learned from course to use	15	88%
Participants who feel their learners have benefitted from course	16	94%

Standard 4. Knowledge and skills which participants acquire through this program will benefit both the program and the low-literate learner.

Evaluators concluded that both the program participants and

their low-literate learners were able to benefit from the knowledge and skills acquired through the program. Participant responses to items on the follow-up questionnaire which relate to this standard indicate strong agreement that both the participants and the low-literate learner derived benefit from the course (see Table 13). Support for this conclusion is also indicated by participant responses to open-ended questions posed in the follow-up questionnaire; there was a general consensus among participants that the course made their tutoring sessions more interesting and more productive for their learners (see Descriptive Data, p. 136).

Table 13

Participants' rating of course

	4 Strongly agree 3 Agree 2 Disagree 1 Strongly disagree	
The literacy providers course:	N	Rating
Gave me information I could use	28	3.7
Helped my students as well as me	17	3.4
Could have contributed more to my practical knowledge	28	2.6
Could have provided me with more theory	28	2.2
Made me a better tutor	28	3.5
Gave me more confidence as a literacy provider	28	3.9
Should be offered again for others	28	3.9
Can benefit literacy providers with a variety of levels of experience	28	3.2

Descriptive Data

Data from open response items of the follow-up questionnaire are summarized as follows.

Items 5.1 and 5.2 asked respondents if participation in the Literacy Providers' course helped them as tutors. While not all

respondents had the opportunity to tutor within the six months after the completion of the course, all 28 responded positively to these items. Those who had not yet had the opportunity to apply their new skills and knowledge nevertheless felt confident that their tutoring ability would be improved as a result of participation in the course.

Respondents noted the following specific improvements in their tutoring abilities:

- Use of alternative instructional methods - i.e., language experience.
- Self-assurance that techniques being used are of value.
- Access to and use of new tutoring materials and resource people.
- Use of new techniques for tutoring approaches in spelling.
- Greater awareness of tutoring in relation to individual learner needs.
- Knowledge of stages of literacy and their relationship to appropriate tutoring.
- More effective approaches in tutoring of spelling.

Sample verbatim comments from respondents, extracted from the follow-up questionnaire, who had tutored since the completion of the course include:

Respondent # 6

I learned a lot of new techniques; new approaches for individuals - adults. I got a lot of information that I could use.

Respondent # 11

It gave me better training. I felt I didn't have any real skills before.

Respondent # 14

The spelling sessions and whole language experience sessions were particularly helpful. It was exactly what I needed. My student and I are much more effective in our sessions together.

Respondent # 19

It has given me a lot of new ideas and techniques, and it was nice to have feedback from other tutors. The new resource materials were good too.

Respondent # 25

It gave me a more global picture. It gave me more approaches - a broader perspective as a tutor, which is what I do mostly.

Respondent # 28

I wish I had the course before I started in the area. I had to do a lot scratching.

Sample verbatim comments from respondents who had not tutored since the completion of the course include:

Respondent # 7
It broadened my scope. Prior to the course I was always looking for ways of teaching and it helped me a lot.

Respondent # 8
It spelled out a lot of things for me, especially about the students - how to approach the student and how to get the best response from the student.

Respondent # 16
It kept me on top of what the tutors are doing and what their needs are. It refreshed me.

Respondent # 22
It gave me more insight. I got suggestions from the other participants and some good ideas from the speakers.

Item 6 from the follow-up questionnaire asked respondents to provide examples of actual ideas, approaches, or methods that they had implemented since the completion of the course. All respondents who had the opportunity to tutor since the completion of the course provided specific examples that they had implemented, which are summarized below:

- CLOZE procedure.
- Language experience approach.
- Development of personal dictionaries.
- The use of everyday things as teaching tools, rather than just book s.
- The use of materials suited to adults.
- Story creation techniques.
- The use of student diaries.
- The use of environmental materials.

Item 7 from the follow-up questionnaire asked respondents to consider which specific parts of the Literacy Providers' course contributed most to their practical, day-to-day tutoring work. Specific parts of the course that contributed significantly to tutoring experiences were for the most part the same as those which respondents indicated were implemented by them in their tutoring sessions: They include:

- Specific techniques and methods in language experience.
- Specific information on stages of literacy.
- The interaction between those taking the course and guest presenters.
- Information on the readability formula.
- Specific information on reading disabilities.
- Specific information on spelling techniques.

In responding to this item, many of those interviewed pointed to other positive attributes of the course which might not lead to direct implementation, but which nonetheless contributed to the respondents' growth and development, and hence to their abilities as literacy providers. The most commonly mentioned positive attribute was the sharing of experiences with presenters and with other literacy providers. The resource manual and the videos also were deemed of great value, and to be of general help in tutoring work. Sample verbatim comments are as follows:

Respondent # 1
Just talking to other people and hearing what others had to say was helpful.

Respondent # 9
The first couple of lessons concerning characteristics of adult learners and tutors...steered me in the right direction with regard to tutoring adults.

Respondent # 12
I enjoyed the sessions where the low-literate individuals spoke.

Respondent # 21
It all contributed - hearing other people's ideas.

Respondent # 26

Listening to other people's experiences, particularly those working in the field at the formal level. These people had much to contribute.

Respondent # 28

I found the resource manual very helpful - and the videos. Just being able to discuss ideas and problems with others in the field was very helpful.

Item 8 from the follow-up questionnaire asked respondents to identify those parts of the Literacy Providers' course that they felt contributed little or nothing to their practical, day-to-day tutoring work. Specific parts of the course that did not contribute to their tutoring experiences were very few. Approximately half of the respondents mentioned that they found all parts of the course to be of practical value, and they enjoyed participating in all components of the course. Ten respondents noted specific parts of the course that made little contribution to their own growth and development as tutors, or to their tutoring practices. The most frequently mentioned part of the course that made little contribution was the session on learning disabilities. Five respondents felt that the topic was important, but that coverage was too theoretical, and lacking in practical application information.

Another part of the course that was criticized as irrelevant

was the open discussion, which was perceived as going on too long, rambling, and failing to relate helpful experiences. It should be noted, however, that this is a minority perception. Approximately 80% of participants in the course enjoyed the interaction among tutors and found it very beneficial.

Items 9 and 9.1 asked respondents to describe how those being tutored benefitted from the literacy providers' participation in the course. The most commonly mentioned benefits were:

1) the implementation of a greater variety of methods, approaches and materials;

2) improved self-confidence on the part of the tutor which led to the willingness to try new approaches and materials.

Specific verbatim comments of respondents are as follows:

Respondent # 5
I have more to offer my students. Previously all I had was 10 hours of Laubach training. I now have more specific techniques that I can use in given situations.

Respondent # 6

The course gave me the confidence to experiment, plus more sources of new information and more things for her (the low-literacy learner) to do.

Respondent # 12

I have more confidence in myself. The learners have problems I hadn't really thought about before.

Respondent # 14

There is a definite improvement in my approach and my learners are not depending on me so much.

Respondent # 18

The sessions with my students are more productive and interesting - less repetitious and boring. I have used the First Time Reader supplied by the St. John's participant.

Respondent # 25

The case studies improved my sensitivity. I learned to be careful not to treat low-literacy people as if they were in Kindergarten. I have become very aware of the importance of this idea.

Respondent # 26

I have more to bring them so I'm sure they have a better feeling

about our sessions together. I felt better about tutoring so I think that translates into a positive outcome.

CHAPTER 5

Summary, Conclusions and Recommendations

Summary

The researcher chose the Stake Responsive Evaluation approach, a variant of what House (1978) describes as a transaction model, to evaluate the DELP Course. The suitability of this approach was established through a review of numerous evaluation models, and the successful implementation of the approach by Lertpradist (1990). The Responsive Evaluation approach was chosen because: (a) its emergent and naturalistic methods allow the evaluator to be both flexible and sensitive to programs in which the social setting and context plays an integral role; (b) it is highly democratic in the way it addresses the information needs of all audience groups; and (c) it provides results which are oriented to providing meaningful information to a diverse group of people.

Recommendations

This section contains recommendations for the future implementation of the Distance Education Course for Literacy Providers. The recommendations are organized around five evaluation questions which were posed by evaluators prior to the beginning of data collection. They are based on the judgements made by evaluators in relation to the standards and on the

conclusions presented in the previous section.

Question 1. In terms of program rationale, is the program as designed and implemented, a viable program?

Evaluators believe that the program is viable from a rationale perspective. Originally designed to help those involved in literacy provision by providing training that would (1) meet the information/knowledge needs of both trained and untrained literacy providers, and (2) enable literacy providers to improve their literacy provision, data indicate that the program did meet the needs of all participants.

The program as implemented was viable. Participants indicated that the program timing, program length, teleconference length, teleconference frequency, and program support materials were all satisfactory, and evaluators observed that for the most part the program as implemented ran smoothly.

Recommendation 1. That the program, in future offerings, be of ten weeks duration.

Recommendation 2. That the program, in future offerings, be offered once weekly for approximately two hours per week.

Recommendation 3. That the program, in future offerings, be available to both groups of literacy providers, namely, those with previous training and those with no previous training.

Question 2. In terms of impacts/effects, did participants increase their knowledge and competency in the area of adult literacy?

Evaluators deemed that participants increased their knowledge and competency in the area of adult literacy. All participants indicated that their knowledge and understanding of adult literacy was enhanced by attending the course, and all participants indicated that they would recommend the course to other literacy providers, should it be offered again. This question could only be addressed through participant testimony, since there was no pre-program measure of participants' knowledge/competency levels. In future offerings it would be ideal to establish participants' entry level knowledge and competency for comparison purposes.

Recommendation 4. That future offerings of the program include participant pre-questionnaires regarding their entry level knowledge and competency.

Question 3. In terms of impacts/effects, did participants experience a change in attitude toward the positive regarding adult literacy?

Evaluators believe that participants' attitudes toward their work in literacy provision have been enhanced by completing the course. While participants, most of whom were volunteer literacy providers, obviously felt positive about the literacy issue prior to enroling in the course, they did not necessarily feel that they were making a significant contribution. Many of the participants were unsure of their effectiveness as literacy tutors, and they felt the need for confirmation and reassurance that they were using appropriate strategies and tactics. The course provided that reassurance, and all participants, at the end of the course, felt that they could return to literacy provision with something worthwhile to offer.

Question 4. In terms of program objectives, did the program result in the attainment of all objectives?

Evaluators deemed that all participants achieved those objectives that were important to them as individuals. In a course such as this it is not important that all participants achieve the same objectives, and indeed given the varying backgrounds of participants it should not be expected that they want the same things from the course. Evaluators did ascertain,

through their data, that the general goal statements guiding each teleconference were achieved by all participants.

Recommendation 5. That in future offerings of the program participants be asked to set their own objectives prior to the beginning of the course, so that those objectives, in conjunction with the general program goal statements, can be assessed by evaluators.

Question 5. In terms of program design, was the Distance Education Course for Literacy Providers the most efficient and effective vehicle for achieving program objectives?

Evaluators deemed that the course as designed was an effective vehicle in that program objectives were achieved. In terms of efficiency, the program as designed and implemented had the ability to involve participants in various locations across the province, and to expose these distant participants to various knowledgeable and interesting guest presenters. While there is considerable cost involved in using the audio teleconference delivery system, that cost is considerably less than bringing participants and presenters together in a particular geographic location.

Evaluators are cognizant of the fact that the course, with

little further development, could be offered just as efficiently and effectively through other delivery modes, however. Given that literacy providers are mainly volunteers with other daily and weekly duties, responsibilities, and work commitments, it is not feasible or cost-effective, for the most part, to offer the course via live instruction in selected central locations around the province. However evaluators are cognizant of the fact that the course, with little further development work, could be fully packaged so that a local resource person or literacy council could coordinate offerings for a given area. All that would be required would be to develop the teleconference sessions as a series of audiotapes or videotapes. Participant interaction and project reports would take place within the local group, and in fact, the course would have portability beyond the boundaries of the province.

Another possibility for future offerings of the program would be to format it as a two day intensive workshop to be held in various locations over weekends. Given the administrative time involved in the teleconference sessions, that aspect of the course could be reduced by four to five hours, making it possible to conduct the course as a two day workshop. And if the various topics within the course were modularized, a series of half-day, or one day workshops could be offered. Evaluators feel that the program has the potential, with a little more development work, to be offered in a variety of formats through a variety of modes, adding to the efficiency and effectiveness of providing training

for literacy providers around the province and beyond.

Recommendation 6. That the Distance Education Course for Literacy Providers be developed as fully packaged modules so that it can be offered, in the future, in a variety of formats through a variety of modes.

Conclusion

This study set out to examine the various models for program evaluation as described by House (1978) in his Taxonomy of Major Evaluation Models with a view toward identifying a methodology appropriate for application to any small to medium scale community-based distance education programs for semi-professional development. The results of the study affirms that a transactional evaluation model represents an effective, efficient, rigorous, and socially appropriate methodology for evaluating such programs. It also supports the opinion of Lertpradist (1990) who used Robert Stake's responsive evaluation approach to evaluate an artificial fish breeding program in rural Thailand.

The responsive evaluation approach was designed to emphasize evaluation issues that are important for a particular program (Stake, 1983). Unlike more preordinate models of evaluation, the

responsive approach does not assume that the only evaluation issues that are important are those that derive from a single, and thereby elite evaluation audience. Rather, it is very democratic in that it solicits the concerns and issues of all stakeholding audiences associated with a program, and measures program outcomes in relation to them.

Democracy and participative management are important considerations in the evaluation of programs which are intended to support community-based economic, educational, and social development activities. The Distance Education Course for Literacy Providers was a program intended to provide community-based workers, both volunteer and professional, with skills and tools with which to reduce the level of adult illiteracy within their communities. The larger goal of the program was to emancipate and empower a generation of illiterate Newfoundlanders so that they may participate, develop, and attain a reasonable degree of financial security in a society that requires literacy.

Stakes responsive evaluation approach provided those individuals who are closest to the front line of the problems being attacked by programs such as the Distance Education Course for Literacy Providers with a sense that they are full and significant players in the direction of their destiny. It provided them with a sense of control and ownership of their problems and their resolutions. In sum, it encouraged, recognized, and respected self-determination by placing value on, and responding to, the needs of the audiences to whom a given

program is addressed.

References

Alvarado, J., D'Agostino, G., & Bolanos, G. (1991). Orientation and conceptualization of the academic quality control centre. In R. Schuemer (Ed.), Evaluation concepts and practices in selected distance education institutions. FernUniversitat, Hagen, West Germany. (ERIC Document Reproduction No. 329 784).

Batey, A. & Cowell, R.N. (1986). Distance education: an overview. Washington, D.C.: Office of Educational Research and Improvement. (ERIC Document reproduction No. 278 519)

Berk, R.A. & Rossi, P.H. (1991). Thinking about program aluation. Newbury Park, CA: Sage Publications.

Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive Domain. New York, NY: David McKay.

Borg, W.R. & Gall, M.D. (1989). Educational research: An introduction (5th ed.). New York: Longman.

Bookfield, S.D. (1986). Understanding and facilitating adult learning. San Francisco: Jossey-Bass Publishers.

Caplan, N., Choy, M.H. & Whitmore, J.K. (1992). Indochinese refugee families and academic achievement. Scientific American, 266(2), 36-42.

Cippola, C.M. (1978). The economic history of world population (7th ed.). New York, NY: Penguin Books.

Committee on Evaluation, Phi Delta Kappa National Study. (1971). Educational evaluation and decision making. Itasca, IL: F.E. Peacock.

Cross, A.F. (1992). Development and implementation of a program of evaluation to support improvement of organizational components in a child care centre. Doctoral Practicum Paper, Nova University. (ERIC Document Reproduction Service No. ED 343 671)

Dillon, C. & Gunawardena, C., (1992). Evaluation research in distance education. British Journal of Educational Technology, 23(3), 181-194.

Donald, J. (1983). How illiteracy became a problem (and literacy stopped being one). Journal of Education, 165(1), 35-52.

Downs, R.B. (1974). Horace Mann: Champion of public schools. New York: Twayne Publishers.

Elsey, B. (1993). Voluntaryism and adult education as civil society and the 'third way' for personal empowerment and social change. International Journal of Lifelong Education, 12(1), 3-16.

Evans, T. & Nation, D. (1990). Introduction. In T. Evans & D. Nation (Eds.), Critical reflections on distance education. London: The Falmer Press.

Fetterman, D.M. (1986). Conceptual crossroads: methods and ethics in ethnographic evaluation. In D.D. Williams (Ed.), Naturalistic Evaluation. San Francisco, CA: Jossey-Bass, Inc..

Fiorino, A.F. (1978). The moral education of Egerton Ryerson's idea of education. In N. McDonald and A. Chaiton (Eds.), Egerton Ryerson and his times. Toronto: Macmillan.

Foner, E. (1988). Introduction. In Paine, T. (1791,1792/1988) The Rights of Man. New York, NY: Penguin Books.

Fox, S. (1991). The production and distribution of knowledge through open and distance learning. In D. Hlynka and J.C. Belland (Eds.), Paradigms Regained: The uses of illuminative, semiotic and post-modern criticism as modes of inquiry in educational technology (pp.217-239). Englewood Cliffs, NJ: Educational Technology Publications.

Franklin, J.L. & Trasher, J.H. (1976). An introduction to program evaluation. New York: Wiley.

Gagne, R.M., Briggs, L.J., and Wager, W.W. (1988). Principles of instructional design (3rd ed.). New York, NY: Holt, Rinehart & Winston.

Gagne, R.M. (1989). Studies of learning: 50 years of research. Tallahassee, FL: Florida State University.

Garrison, D.R. (1986). Researching dropout in distance education. Distance Education, 8(1), 95-101.

Gough, H. (1988). The legitimization of the radical tradition in France, 1789-1901. The International Journal of Social Education, 3(1), (9-20).

Goodson, I. (1988). Curriculum as structured inequality. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED 305 737).

Guba, E.G. & Lincoln, Y.S. (1981). Effective evaluation. San Francisco, CA: Jossey-Bass, Inc..

Guba, E.G. & Lincoln, Y.S. (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. In D.D. Williams (Ed.), Naturalistic Evaluation. New Directions for Program Evaluation, no. 30. San Francisco, CA: Jossey-Bass.

Guba, E.G. & Lincoln, Y.S. (1989). Fourth generation evaluation. Newbury Park, CA: Sage Publications.

Guillemet, P., Bedard, R., & Landry, F. (1986). Tele-University du Quebec. In D. Kaufman and I. Mugridge (Eds.), Distance education Canada (pp.144-160). Dover, NH: Croon Helm.

Handcock, W.G. (1989). So longe as there comes noe women: Origins of English settlement in Newfoundland. St.John's, NF: Breakwater Books.

Harry, K. (1990). The Open University, United Kingdom. In Koul, B.N. & Jenkins, J. (Eds.)., Distance education: A spectrum of case studies. London: Kogan Page.

Hess, R.D. & Azuma, H. (1991). Cultural support for schooling: contrasts between Japan and the United States. Educational Researcher, 20(9), 2-8,12.

Holmberg, B. (1986). Growth and structure of distance education. London: Croon Helm.

House, E.R. (1978). Assumptions underlying evaluation models. Educational Researcher, 7, 4-12.

House, E.R. (1980). Evaluating with validity. Beverly Hills: Sage Publications.

House, E.R. (1986). Evaluation and legitimacy. In E.R. House (Ed.), New directions in educational evaluation (pp.5-9). London: The Falmer Press.

Hope, C. (1986). Private correspondence schools. In I. Mugridge & D. Kaufman (Eds.), Distance Education in Canada (pp. 204-213). Doyer, NH: Croon Helm.

Hornstein, G.A. (1992). The return of the repressed: Psychology's problematic relations with Psychoanalysis. American Psychologist, 47(2), 264-263.

Innis, H.A. (1951). The Bias of Communications. Toronto, Ont: University of Toronto Press.

Innes, A.H. (1982). Biases: Threats to validity in evaluation models. Paper presented at the Annual Meeting of the American Educational Research Association, New York: NY. (ERIC Document Reproduction Services No. ED 218 352).

Keegan, D. (1990). Foundations of distance education (2nd ed.). London: Routledge.

Kirkup, G. (1986). The feminist evaluator. In E.R. House (Ed.), New directions in educational evaluation. London: The Falmer Press

Knapper, C., et. al. (1985). Instructional development for distance education. Paper presented at the World Conference of the International Council for Distance Education, Melbourne: Australia (ERIC Document Reproduction Services No. ED 265 638).

Leach, E.E. (1983). Tuning out education: The cooperation doctrine in radio, 1922-38. (ERIC Document Reproduction No. ED 248 835).

Lertpradist, P. (1990). A study of the application of a selected evaluation methodology in an extension setting. Unpublished Master's thesis, Memorial University of Newfoundland, St, John's, NF.

Lewis, T.R. & Nichols, R.G. (1965). Speaking and listening: A guide to effective oral-aural communication. Dubuque, Iowa: Wm. C. Brown Company Publishers.

Logsdon, D.M., Taylor, N.E., & Blum, I.H. (1988). It was a good experience: the problems and trials of implementing and evaluating a parent participation program. In D.M. Fetterman (Ed.), Qualitative approaches to evaluation in education. New York, NY: Praeger Publishers.

Mager, R. (1962). Preparing instructional objectives. Palo Alto, CA: Fearon Publishers.

McCann, W.P. (1988). Denominational education in the twentieth century in Newfoundland. In W.A. McKim (Ed.), The vexed question (pp. 60-79). St.John's, NF: Breakwater Books.

McInnis-Rankin, E. & Brindley, B. (1986). Student support services. In D. Kaufman. & I. Mugridge (Eds.), Distance Education in Canada. London: Croon Helm.

Nevo, D. (1986). Conceptualization of educational evaluation: An analytical review of the literature. In E.R. House (Ed.), New directions in educational evaluation. London: The Falmer Press.

Nowakowski, J.R. (1983). On educational evaluation: A conversation with Ralph Tyler. Educational Leadership, 40(8), 24-29.

Osborne, J.W. (1991). Humanistic learning and teaching. In R.H. Short, L.L. Stewin & S.J.H. McCann (Eds.), Educational Psychology: Canadian Perspectives. Toronto, Ont: Copp Clark Pitman Ltd.

Owens, T.R. & Wolf, R.L. (1985). Adversary approach to evaluation. In D.L. Stufflebeam and A.L. Shinkfield (Eds.), Systematic evaluation: A self-instructional guide to theory and practice. Boston, MA: Kluwer-Nijhoff Publishing.

Patton, M.Q. (1980). Qualitative evaluation methods. Beverly Hills, CA: Sage Publications.

Patton, M.Q. (1982). Practical evaluation. Newbury Park, CA: Sage Publications.

Perry, W. (1990). The genesis of the first open university. In Koul, B.N. & Jenkins, J. (Eds.)., Distance Education: A Spectrum of Case Studies. London: Kogan Page.

Peters, O. (1983). Distance teaching and industrial production: a comparative interpretation in outline. In Sewart, D., Keegan, D. & Holmberg, B. (Eds.). Distance education: international perspectives. London: Croon Helm.

Popham, W.J. (1988). Educational evaluation (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Provost, M. (1973). Evaluation of ongoing programs in the public school system. In B.R. Worthen & J.R. Sanders (Eds.). Educational evaluation: Theory and practice. Worthington, Ohio: Charles A. Jones Publishing.

Reese, W.L. (1980). Dictionary of philosophy and religion: Eastern and Western thought. Atlantic Highlands, NJ: Humanities Press.

Reigeluth, C. (1983). Instructional design: Theories and models. Hillsdale, NJ: Lawrence Erlbaum Associates.

Reiser, R.A. (1987). Instructional technology: A history. In R. Gagne (Ed.). Instructional Technology: Foundations. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rivlin, A.M. (1971). Systematic thinking for social action. Washington: The Brookings Institution.

Robertson, D.L. (1982). Toward a model for ESL program evaluation. Paper presented at the Annual Convention of Teachers of English to Speakers of Other Languages, Toronto, Ontario. (ERIC Document reproduction No. ED 234 607)

Romizowski, A.J. (1981). Designing instructional systems: Decision making in course planning and curriculum design. London: Kogan Page.

Romiszowski, A.J. (1988). The selection and use of instructional media (2nd ed.). London: Kogan Page.

Rossi, P.H. & Freeman, H.E. (1993). Evaluation: A systematic approach (5th ed.). Newbury Park: Sage Publications.

Rutman, L. (1980). Planning useful evaluations. Evaluability assessment. Beverly Hills, CA: Sage Publications.

Sadler, D.R. (1981). Intuitive data processing as a potential source of bias in naturalistic evaluations. Educational evaluation and policy analysis, 4(3), 25-31.

Salasin, S. (1974). Exploring goal-free evaluation: An interview with Michael Scriven. Evaluation, 2(1), 9-17.

Scriven, M.S. (1987). Prose and cons about Goal-Free evaluation. In B.R. Worthen & J.R. Sanders (Eds.), Educational evaluation: Alternative approaches and practical guidelines. New York, NY: Longman.

Saettler, P. (1968). A history of instructional technology. New York: McGraw-Hill Book Company.

Schuemer, R. (1991). Introduction and Preface. In R. Schuemer (Ed.), Evaluation concepts and practices in selected distance education institutions (pp. 1-36). FernUniversitat, Hagen, West Germany. (ERIC Document Reproduction No. 329 784).

Scriven, M. (1986). Evaluation as a paradigm for educational research. In E.R. House (Ed.), New directions in educational evaluation (pp. 53-67). London: The Falmer Press.

Seaborne, K. & Zuckernick, A. (1986). Course design and development. In D. Kaufman & I. Mugridge (Eds.). Distance Education in Canada. London: Croon Helm.

Seels, B. (1989). The instructional design movement in educational technology. Educational Technology, 29(5), 11-15.

Simon, B. (1960). Studies in the history of education: The two nations and the educational structure 1780-1870. London: Lawrence and Wishart.

Skinner, B.F. (1954). The science of learning and the art of teaching. Harvard Educational Review, 24, 86-97.

Skinner, B.F. (1958). Teaching machines. Science, 128, 969-977.

Smith, L.D. (1992). On prediction and control: B.F. Skinner and the technological ideal of science. American Psychologist, 47(2), 216-223.

Smith, K.C. (1983). External studies at New England: A silver jubilee review, 1955-1979. In D. Sewart, D. Keegan and B. Holmberg (Eds.). Distance education: international perspectives. London: Croon Helm.

Southam Newspaper Group. (1987). Literacy in Canada - A research report. Ottawa: Author.

Stake, R.E. (1975). To evaluate an arts program. In R.E.Stake (Ed.), Evaluating the arts in education. Columbus, OH: Charles E. Merrill Publishing.

Stake, R.E. (1977). Robert Stake: responsive evaluation. In D. Hamilton, B. MacDonald, C. King, D.Jenkins & M.Parlett (Eds.), Beyond the numbers game. Berkeley, CA: McCuthcan Publishing Corp.

Stake, R.E. (1983). Program evaluation: particularly program evaluation. In G.F. Madaus, M.S. Scriven & D.L. Stufflebeam (Eds.), Evaluation models: viewpoints on educational and human services evaluation. Boston, MA: Kluwer-Nijhoff Publishing.

Stahmer, A., & Helm, B. (1987). Communications technology and distance learning in Canada. A survey of Canadian activities. Ottawa, Ont: Department of Communications. (ERIC Document Reproduction No. ED 290 464)

Stubbs, J., et. al. (1985). Organizational diversity among dual-mode institutions. Paper presented at the World Conference of the International Council for Distance Education. Melbourne, Australia. (ERIC Document Reproduction No. 265 640).

Stufflebeam, D.L. & Webster, W.J. (1983). An analysis of alternative approaches to evaluation. In G.F. Madaus, M.S. Scriven & D.L. Stufflebeam (Eds.), Evaluation Models: Viewpoints on educational and human services evaluation (pp.23-43). Boston: Kluwer-Nihoff Publishing.

Stufflebeam, D.L. & Shinkfield, A.J. (1985). Systematic evaluation: A self-instructional guide to theory and practice. Boston: Kluwer-Nijhoff Publishing.

Stufflebeam, D.L. (1987). Comment by Daniel L. Stufflebeam: Should or can evaluation be goal-free? In B.R. Worthen & J.R. Sanders (Eds.), Educational evaluation:Alternative approaches and practical guidelines. New York, NY: Longman.

Sungalia, H. (1980). A new model for program evaluation: The appreciative one!. Victoria, Australia: Deakin University, School of Education. (ERIC Document Reproduction Service No. ED 231 137)

Sweet, R. (1991). Canadian proprietary correspondence schools: some issues of access and technology. Journal of Distance Education, 6(1), 42-63.

Thomas, R.M. (1985). Comparing theories of child development (2nd ed.). Belmont, CA: Wadsworth Publishing Company.

Thompson, M.S. (1975). Evaluation for decision in social programs. Westmead, UK: Saxon House.

Thornton, N. (1990). Women's studies, Australia. In B.N. Koul and J. Jenkins (Eds.). Distance education: a spectrum of case studies. London: Kogan Page.

Toynbee, A. (1967). Lectures on the industrial revolution of the 18th century in England (The industrial revolution defined). In Tierney, B., Kagan, D. & Williams, L. P. (Eds.) The Industrial Revolution in England - Blessing or Curse to the Working Man? New York, NY: Random House.

Turner, D.H. (1979). Behind the myths: An introduction to the structural study of myths and folklore. In Turner, D.H., & Smith, G.A.. (Eds.), Challenging Anthropology: A Critical Introduction to Social and Cultural Anthropology. Toronto: McGraw-Hill Ryerson.

Turner, D.H. & Smith, G.A.(Eds.) (1979). Challenging Anthropology: A Critical Introduction to Social and Cultural Anthropology. Toronto: McGraw-Hill Ryerson.

Tyler, R.W. (1980). A brief overview of program evaluation. Journal of Cooperative Education, 16(3), 7-15.

Tyler, R.W. (1949). Basic principles of curriculum and instruction. Chicago: University of Chicago

Tyack, D.B. (1976). Ways of seeing: an essay on the history of compulsory schooling. Harvard Educational Review, 46(3), 355-389.

Webber, C.F. (1987). Program evaluation: A review and synthesis. (ERIC Document Reproduction Service No. 291 771).

Williams, D.D. (1986). Editor's note. In D.D. Williams (Ed.), Naturalistic Evaluation. San Francisco, CA: Jossey-Bass, Inc..

Worthen, B.R. & Sanders, J.R. (1973). Educational evaluation: Theory and practice. Worthington, OH: Charles A. Jones Publishing.

Worthen, B.R. & Sanders, J.R. (1987). Educational evaluation: alternative approaches and practical guidelines. New York: Longman.

Worthen, B.R. & Sanders, J.R. (1991). The changing face of educational evaluation. Theory Into Practice, 30(1), 3-12.

Verduin, J.R. & Clark, T.A. (1991). Distance education: The foundations of effective practice. San Francisco, CA: Jossey-Bass Publishers.

APPENDIX A

Indepth Interview Data Summary

Participant 1

P. 1 is a former teacher, who had taught in the school system for nine years. She has been involved in literacy provision for approximately six years. She has been teaching literacy courses at a local community college. She is in favour of the two-tiered approach to literacy provision - offered through both community college courses and through the volunteer sector. Prior to participation in this course, she had no literacy training.

She believes that some sort of course should be a requirement before beginning literacy provision. Despite the fact that she had been a teacher some years previously, she found it difficult starting out with no training.

P. 1 believes that the literacy problem in Newfoundland is a serious one. She feels that students are just pushed on through the school system and not given the help they need. She also feels that poor home influences are a factor in the literacy problem, but she recognizes that she doesn't know the causes, and she feels that a significant portion of literacy funding should be focused on investigation of what is causing so much illiteracy in Newfoundland.

P. 1 considered the course a success. She had hoped to build her

confidence, in addition to gaining new knowledge, and she indicated that the course accomplished both of her personal goals. I found out that I was on the right track in what I was doing. It's important to know if what you're doing is right.

P. 1 felt that there were very few things in need of improvement in the course. She considered the greatest area of weakness to be the learning disabilities teleconference. The information presented was too theoretical - it was hard to take any notes, and there wasn't enough practical tips on how to help those with specific disabilities.

P. 1 also felt that there was too much chatter throughout the teleconferences. She felt that the sharing and interaction was overdone, and not directed enough to be of benefit. Lots of the participation was a waste of time and off track.

Overall, P. 1 enjoyed the course and found it very beneficial.

Participant 2

P. 2 is a substitute/replacement teacher who has taught reading in the school system for approximately two decades. Through volunteer work in the community, she became aware of the need for literacy volunteers and the availability of Laubach literacy training courses, so she completed training and became a

volunteer tutor. She estimates that she has had about 250 hours of experience in literacy provision. Her students have ranged from GED level to totally illiterate.

P. 2 thinks that most literacy problems in Newfoundland begin at home - no emphasis on education, poor nutrition, not enough rest, and emotional trauma - you see the beginnings in the kids in school. She thinks that government funding should be aimed at removing the stigma of illiteracy so that those needing help will seek it. She also disagrees with the offering of literacy courses through the community college sector. The government is wasting money to just keep people employed. The volunteers can do the work at the community level to help with the problem - there's no need for paid tutors and courses.

Despite her extensive training and experience, P. 2 found the course beneficial, and she felt that it complemented her Laubach training. Mostly I found that this course confirmed what I already knew. But I did learn a few new things. The best part of the course, for P. 2, was its emphasis on incorporating student writing into literacy provision. She indicated that the program met her needs, and she particularly appreciated the course manual, which she felt was excellent.

In terms of improvements, she would have liked more on learning disabilities. She didn't think that topic was covered well. In addition she felt that the spelling module, while beneficial,

could have been more effective if participants had presented their questions the week before, and the resource person addressed the questions as part of the presentation. P. 1's major criticism concerned the participant interaction/feedback during the teleconferences. I thought all that asking for feedback was a waste of time - listening to others talk about stuff that wasn't relevant to my needs - I didn't get any new ideas out of it.

P. 2 felt that the course would be of benefit to everyone, and she as a trained and experienced literacy provider benefitted from her involvement. She thinks the course should be offered in the future, given that there is always a need to enhance skills, and that there are still many untrained literacy providers out there.

Participant 3

P. 3 is a former teacher in both the school system and the community college system. He has seventeen years teaching experience. He has also been employed as a community development worker. He became involved in literacy provision through his community development work, and through awareness of the extent of the problem through the efforts of MUN's Extension Service.

P. 3 has completed some training, specifically the Laubach course offered in his local area. He felt the Laubach course was very

practical, and focused on actual things you could do day to day with a literacy client.

P. 3 feels that literacy is a big problem in Newfoundland, although not as big as estimated. He feels that government funding of literacy in the province should go into research for the long term. But in terms of the present he feels that program development is needed. He disagrees with the current two-tiered system of literacy provision through both community colleges and the volunteer sector. I think we are defeating our own ends. There are too many programs and people competing with each other. There is not enough leadership or centralization of the effort. While he praises the work of volunteers (being one himself), he doesn't think that they alone can deal with the issue. You need some group to administer and coordinate things.

P. 3 did not think that the Literacy Providers' Course was as beneficial as Laubach training. He had assumed that the course would only be a couple of weeks long, and for such a long course he didn't feel that it offered much of practical help to literacy providers, particularly those new to the field. You don't learn much practical - it seems to be just an extension of someone's theory. He did feel that the course manual was good, and that he got enough out of the course to be able to help other tutors. He found the sessions on spelling and language experiences helpful.

P. 3 thought that there was plenty of time for participant interaction/sharing - too much, in fact. There was nothing wrong with that (listening to participants) in itself, but there was too much time spent on it for most people. He though the course should be offered for those with considerable experience, and that more basic stuff was needed for new tutors.

Participant 4

P. 4 was trained as a primary/elementary school teacher. For the past few years, she has been employed as a substitute teacher, teaching all levels and grades. She became involved in volunteer literacy provision four years ago, and completed Laubach training of approximately 10 hours. Since then she has only tutored one person. She has worked as a local literacy coordinator for the past few years.

P. 4 says that there are now about thirty trained tutors in her area of the province, but that few people come forward for tutoring. She thinks that government funding should focus on publicizing and promoting literacy, and on what's available in the community to help those who have low literacy levels. She says that government only recently began to recognize how big a problem illiteracy is in the province. She feels that volunteers do a tremendous job of literacy provision, but she agrees with the idea of everyone being involved. She supports the community

college courses as well.

P. 4 felt that the Laubach focus on the phonics approach was too narrow. Her Laubach training did not provide her with the knowledge to deal with the many types of problems she met when she began tutoring. She felt that this course had greater application. This course is much broader and much more informative. P. 4 particularly liked the project she completed; locating and adapting suitable curriculum materials. She feels that one of the biggest problems that literacy tutors have to contend with is finding suitable reading materials for their particular clients.

P. 4 thought the course was suitable for everyone involved in literacy provision. She particularly valued the manual, which she saw as serving the purpose of a refresher course for her, plus having relevant things for other tutors in the area. I don't think the course could have been better. It touched on just about everything. By doing the course I can pass on the information to tutors.

In terms of improvements, P. 4 commented on a number of areas which could be improved for future offerings. She felt that the amount of interaction and commenting by participants was too much, and inclined to waste time. The voicing of opinions by participants was a bit of a waste of time. I think it got out of

hand in some of the teleconferences. I think the comments and conversation by some of the more knowledgeable participants was intimidating for some participants who were new to the literacy area.

P. 4 also noted that they had problems with the electronic system and that most of them at her site were intimidated by it. They were left to set it up themselves at the community college site, and if it hadn't been for one more knowledgeable participant they would have been lost. She thought that an explanation of the system and directions for setting it up should be included in the course materials.

In general she felt that the teleconference component of the course needed improvement. I think if you missed a teleconference you wouldn't have missed a lot. A lot of the teleconferences were just a repetition of what was in the manual. P. 4 felt that the course would have been just as effective without the teleconferences.

Participant 5

P. 5 has worked with MUN Extension for the past seven years, and it was through that organization that she became involved in literacy provision about four years ago. Her training was done through Laubach. She has both tutored and acted as an organizer

of literacy provision in her local area. She found that Laubach training was very focused, while this course provided a more generalized view of literacy. It broadened the horizons of the things we are able to try.

P. 5 thinks that government funding of literacy provision should focus on the school system - as far back as necessary. If you lose a child in the 5th or 6th grade you've lost him. I think the money should be spent in the primary grades. P. 5 doesn't think the community college system of literacy provision is working. She noted You don't need a university degree to be a literacy tutor. There is no need in expanding literacy programs at the community college level if you don't have learners. And a lot of learners won't go to a community college.

P. 5 noted that she had gained a lot from the course. The manual was especially helpful, and she could see
how it would be invaluable in her tutoring situation. She felt that some of the teleconference sessions lacked
good concrete information. She felt that sharing of information from the work/activity sheets didn't add very much to the course.

P. 5's specific area of interest was in learning disabilities. She thought that videotape was very beneficial. P. 5 thought that the course was worthwhile, and provided a good overview of literacy provision, suitable for all involved in the area.

Participant 6

P. 6 has worked with a department of government for a number of years. Through contacts at work she heard about a Laubach seminar that was to be held in the local area, and she decided to attend. Since then she has worked to further the work of the local literacy council and she has tutored one young adult.

P. 6 thinks that the provision of literacy tutoring by local volunteers is very important to improving literacy levels in the province. She thinks that government funding should be focused on preventive measures, such as identifying at-risk groups. She also thinks that more should be spent on providing training for tutors. She doesn't think that the community college system alone is the answer in terms of literacy provision, and she finds it difficult at times that local volunteer tutors are dependent on the community colleges for provision of materials.

P. 6 found that Laubach training was very narrowly focused - suggesting only one approach which she feels does not work well in all situations. She liked the broader view provided by this course, and the focus on the literacy client as a person. She felt that the opportunity to consider and discuss real problems was especially beneficial. Hearing about the problems of others and how they handled them was good and reassuring.

P. 6 had a special interest in learning disabilities prior to the course, but she was disappointed by that session. It was not what I expected. I didn't want information on brain dysfunctions - I wanted to know how to motivate them and find ways to help them overcome their problems. She noted that the audio portion of the videotape was not good - the verbal message was too low and the music too loud. In all that particular session was too theoretical and technical, and not practical enough to help with classroom work.

Despite negative feelings about the learning disabilities session, P. 6 learned a lot of techniques that were very helpful, and felt that she could help other literacy tutors as a result.

Participant 7

P. 7 is a relative newcomer to literacy provision. She knew people in the local area who were involved in literacy provision, and when contacted by them to help out last year she did so, taking over one student. She has completed a Laubach course.

She feel that illiteracy is a serious problem in Newfoundland, and that there is a need for more basic literacy programs because there are so many people out there in need of help. She thinks that community colleges have a significant role to play in literacy provision.

P. 7 noted that this course was very interesting and very helpful. She found the manual excellent. Her Laubach training was very basic and provided a good start, but was not general enough to provide her with knowledge of different methods and approaches in the teaching of reading. She found all of the topics of interest, and she appreciated the practical tips she got from hearing how others dealt with their problems.

P. 7 particularly enjoyed the inclusion of information on learning disabilities, because she has a learning disabled child and could relate to the session quite well. She had no suggestions for improvement in the course.

Participant 8

P. 8 has been a teacher in the regular school system and is now a Laubach teacher. She has completed a lot of training, and she also is experienced in tutoring. She also works in organizing literacy councils in various communities around her local area.

P. 8 thinks that illiteracy is a large and serious problem in Newfoundland. She herself has seen what she refers to as the cycle of illiteracy - moving from one generation to the next in homes where there are no reading materials; where nobody is ever seen reading. Very often children from these homes have trouble with reading in school - and the fact that they have failed in

school prevents them, as adults, from seeking help through community colleges which they identify as schooling.

P. 8 thinks that government funding of literacy should end up directly benefitting the illiterate population of the province. She said that there is enough review, statistics, surveys, and the like - it is time to spend money where it is needed - on training programs for literacy providers, on basic materials for literacy councils, and on literacy programs within the school system.

P. 8 noted that the Laubach training was very basic and very focused on direct and specific interventions to teach reading to adults. She felt that this course was complementary to what she had learned through Laubach. Laubach is basic. This course built on those basics. She felt that she has gained in knowledge and understanding of literacy, and that her existing knowledge was reinforced through the course.

P. 8 found that most of the course was done very well. She enjoyed the teleconferences, and thought that the course manual was excellent as a resource that she could keep and use in her literacy work. She liked the information on stages of literacy, and thought that it was important that all literacy providers be knowledgeable about that topic.

P. 8's only criticism of the course was in the area of learning disabilities. She felt that the information presented was too theoretical, and had little application to her needs in helping learning disabled non-readers. I was looking for basic things. I wanted to direct tutors to resources and send them materials that would help if they had to deal with learning disabilities. She also felt, particularly in relation t the learning disabilities session, that there was too much discussion and too much interaction, particularly from one group of participants. That teleconference lacked organization. We tuned out for a lot of it.

P. 8 felt that overall the course was worthwhile, and that she gained knowledge that she could put to practical use in her tutoring.

APPENDIX B

Evaluation Instruments

Curriculum Committee Guide

Mr./Mrs./Ms./Dr._____, my name is Blair Kettle and I am a graduate student in the Division of Learning Resources of Memorial University's Faculty of Education. As part of my thesis I am working with Dr. Mary Kennedy of the Faculty on an evaluation of the *Distance Education Programme for Literacy Providers*.

You are listed as one of the Curriculum Committe members/ Resource People for that program and I would like to ask you a few questions regarding your involvement. Is this a convenient time?

IF YES: ask question 1.

IF NO: arrange a convenient time. _____

---

1. What was your level of involvement with this programme?

2. What do you feel should be achieved by this programme?

3. What would you say would be a mark of success for the programme?

4. How do you feel about using the distance mode of delivery for the objectives and course content of this programme?

5. Is there anything in particular, other than what you have already mentioned, that you feel should be addressed by this evaluation?

QUESTIONNAIRE FOR LITERACY PROVIDERS
DISTANCE EDUCATION COURSE FOR LITERACY PROVIDERS

Mr./Ms./Mrs._____

My name is _____ and I am calling on behalf of a program evaluation team, headed by Dr. Mary Kennedy of Memorial University, which will be evaluating the *Distance Education Programme For Literacy Providers*. You're listed as one of the intended participants of that programme.

As part of the evaluation process I would like to ask you a few questions about your background in literacy tutoring and the types of things you would hope to learn from the Programme For Literacy Providers.

I will need about 10-15 minutes of your time. Is this a convenient time for you?

IF YES *Proceed to question 1.*

IF NO Could we arrange a convenient time? _____

---

1. Have you had any training as an adult literacy provider? (This includes any sort of formal or informal training).

IF YES go to question 1.1

IF NO go to question 2.1

1.1 What type of agency or group provided your training?

1.2 When did you receive your training?

1.3 In terms of hours, days, weeks or months, how much train' ; have you received?

1.4 If you follow a model or method for literacy tutoring, which model or method do you use? (e.g., Laubach)

1.5 Would you describe for me the area(s) of literacy tutoring in which you feel competent (i.e., good at).

1.6 Which area(s) of literacy tutoring do you hope to gain extra knowledge or competence in? Why?

1.7 What is the least you would hope to get out of your experience with the *Distance Education Programme For Literacy Providers?*

1.8 In your opinion, what should the program accomplish to be considered a success?

1.9 If you were to receive a report of the evaluation, what kinds of information would you like to see included?

1.10 Is there anything else you would like to tell me about the Distance Education Programme For Literacy Providers that we have not already discussed?

Thank you for your time. If you have no objections someone may be getting back to you at another point in the evaluation process.

2.1 Have you ever made efforts to look into adult literacy tutoring methods or techniques? (e.g., through reading or speaking with someone who is knowledgeable about the field). Explain.

2.2 How did you become involved in literacy tutoring?

2.3 Is there a minimum you would hope to learn or achieve from the Distance Education Programme For Literacy Providers? Explain.

2.4 In your opinion, what should the program accomplish to be considered a success?

2.5 If you were to receive a report of the evaluation, what kinds of information would you like to see included?

2.6 Is there anything else you would like to tell me about the Distance Education Programme For Literacy Providers that we have not already discussed?

Thank you for your time. If you have no objections someone may be getting back to you at another point in the evaluation process.

RECORD OF OBSERVATIONS
DISTANCE EDUCATION PROGRAMME FOR LITERACY PROVIDERS

DATE:

TIME BEGIN: TIME END:

TELECONFERENCE LEADER:

ATTENDANCE:

GENERAL NOTES:

STANDARD : (a) - THE CURRICULUM FOR THIS PROGRAM SHOULD SATISFY STUDENT NEEDS.

Criteria: 3. The curriculum has the potential to achieve stated program goals and objectives.

(Make notes on: i. proportions of time dedicated to each teleconference objective, ii. frequency [e.g., Hi Med - Lo] of participant questions per objective, iii. verbalized dissatisfaction with length of time dedicated to, quantity/quality of information provided for, each stated objective)

Criteria: 5. Program participants find the curriculum has sufficient utility to warrant their completion of the program.

(Make notes on: i. participant verbalizations about quitting program because program content is too elementary, ii. general observations about attendance and participant involvement)

STANDARD (a) - PROGRAM PROVIDES OPPORTUNITY FOR SUFFICIENT PARTICIPANT DISCUSSION AND THE SHARING OF IDEAS.

Criteria: 1. Program has regular time scheduled in each teleconference session for participant interaction.
(Make note of the quantity of time allocated for participant discussion)

Criteria: 2. Program includes activities designed to encourage and facilitate participant discussion.
(Make note of types of activities and if they both encourage and facilitate participant discussion)

Criteria: 3. Teleconference moderators/leaders should orchestrate activities which promote participant questions and discussion.

(Make note of efforts on part of moderator to promote participant questions and discussion around stated objective topic. Make note also of quantity of unrelated or tangential discussion)

Distance Education for Literacy Providers Evaluation

Participant Program Completion Questionnaire

Previous Tutoring Experience Yes _____ No _____

PART ONE

Please complete the following items. Circle the yes/no answers, and fill in the blanks where appropriate.

1.	The content of the literacy training program was what I needed.	Yes	No
2.	I was happy with the coverage of the various topics.	Yes	No
3.	I would have liked to have spent more time on certain topic(s).	Yes	No
4.	If yes to #3, list topic(s).		

_____

_____

Answer either 5 or 6

5.	As a person with previous literacy training, the program was helpful to me.	Yes	No
6.	As a person with no previous literacy training, the program was helpful to me.	Yes	No
7.	I really wanted to complete the program.	Yes	No
8.	I stayed with the program mostly because it was something I had started.	Yes	No

Answer either 9, 10, or 11

9.	I attended all of the teleconferences.	Yes	No
10.	I missed some of the teleconferences because I wasn't motivated to attend.	Yes	No
11.	I missed some of the teleconferences because of circumstances beyond my control.	Yes	No
12.	I feel I have really benefitted from participating in the program.	Yes	No
13.	I read the program objectives which were included in my objectives.	Yes	No
14.	The program objectives seemed to be worth achieving.	Yes	No

Answer either 15 or 16

15.	I feel that I have a'tained some of the objectives.	Yes	No
16.	I feel that I have attained most of the objectives.	Yes	No

✓

17. List the specific knowledge areas or topics that you gained, which will be of use to you in your role as a literacy provider.

_____

_____

_____

18. I liked the opportunity for discussion and participation during the teleconferences. Yes No

19. I felt that the presenters encouraged discussion and participation. Yes No

J. I feel that it is important to learn from each other. Yes No

Answer either 21, 22, or 23

21. There was not enough opportunity for discussion. Yes No

22. There was enough opportunity for discussion. Yes No

23. There was too much opportunity for discussion. Yes No

24. Hearing others discuss their insights and problems was really valuable for me. Yes No

25. I believe that I will be able to be more effective in my tutoring because I took part in the program. Yes No

26. The presenters made me feel that I was a valuable part of the program. Yes No

Answer either 27, 28, or 29.

27. The content of the program was too advanced for my needs. Yes No

28. The content of the program was too basic for my needs. Yes No

29. The content of the program was just right for my needs. Yes No

30. I enjoyed the method of presentation for most of the teleconferences. Yes No

31. The teleconference presentations had a good balance of lecture and activities/discussion. Yes No

32. I could have used more practical tips and less theory. Yes No

33. I found the print textbook easy to follow and pleasant to use. Yes No

34. The print materials were attractive and professional looking. Yes No

35. The videotapes were
    (a) interesting Yes No
    (b) relevant to the course Yes No
    (c) easy to understand Yes No
    (d) important to the course Yes No

36.	The teleconference site was easily accessible to me.	Yes	No
37.	The teleconference system was easy to use, once I got accustomed to it.	Yes	No
38.	I had occasional technical trouble with the teleconference system.	Yes	No

If yes to # 38, answer # 39

39.	Because of my problems with the teleconference system, I did not enjoy the course as much as I had hoped.	Yes	No
40.	I found the length of the teleconferences (2 hours) suitable for effective learning.	Yes	No

Answer either 41, 42, or 43.

41.	The use of the teleconference sessions once a week was just right.	Yes	No
42.	Teleconference sessions should have more frequent.	Yes	No
43.	Teleconference sessions should have been less frequent.	Yes	No
44.	I thought the use of teleconferences, videotapes, and print course materials was a good combination of learning experiences.	Yes	No
45.	The course, as delivered, met my learning needs.	Yes	No
46.	I think the literacy providers course is suitable for other types of delivery (i.e., classroom; self-construction package, etc.)	Yes	No
47.	Materials are complete enough so that the course could be delivered without the guest presenters.	Yes	No
48.	I think I would like to deliver parts of the training program in my own area, if materials were accessible.	Yes	No
49.	I would recommend the course to other literacy providers.	Yes	No
50.	I think the course should be offered again.	Yes	No
51.	I would like to do a follow-up course, if one were available.	Yes	No

---

Participant Program Completion Questionnaire

Part Two

52. List the three things that you found most beneficial about the literacy providers' course.

53. List the three things that the literacy providers' training course that you think are most in need of improvement.

54. Complete this sentence. The thing I liked most was

55. Do you feel that this was a good program? Why?

56. If you have anything that you would like yo share with us that hasn't been covered by this questionnaire, please do so here.

Participant Program Rating Scale

Circle the appropriate number to the right

4 strongly agree
3 agree
2 disagree
1 strongly disagree

As a result of my participation in the program.

1.	I have increased my knowledge of literacy issues.	4	3	2	1
2.	I have gained practical tips to help me in my literacy tutoring.	4	3	2	1
3.	I have gained new insights into problems of my learners.	4	3	2	1
4.	I think I will be a better tutor.	4	3	2	1
5.	I feel more able to help my learners.	4	3	2	1
6.	I will not have so many doubts about what I do as a tutor.	4	3	2	1
7.	I think I can apply some of what I learned to my tutoring.	4	3	2	1
8.	I think I can help those of my learners with very difficult problems.	4	3	2	1
9.	I have more confidence as a literacy provider.	4	3	2	1
10.	I feel that I can pass on what I have learned to other literacy providers.	4	3	2	1

Distance Education Course for Literacy Providers
Follow-up Questionnaire

Hello my name is Blair Kettle. As you may recall I am one of the evaluators of the Distance Education Programme For Literacy Providers in which you were a participant last winter. Part of the evaluation design consisted of a follow-up of all the course participants and to that end I would like to ask you a few questions.

NAME: _____

1. Are you currently working as a literacy provider?

IF NO (1.1) Have you worked as a literacy provider since the course ended? _____

IF NO (1.2) Do you intend to continue to work in the field? _____

IF YES (1.3) Do you feel that your participation in the literacy providers course has made you a better tutor? _____

(1.4) Why do you say this?

2. What aspect or aspects of the literacy providers course do you feel contributes most to your tutoring work?

3. What aspect or aspects of the literacy providers course do you feel contributed least to your work as a literacy provider?

4. How, would you say, the people you tutor have benefitted from your participation in the literacy providers course?