

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

EVALUATION OF A SERVICE DELIVERY

PROGRAMME

A thesis presented in partial fulfilment
of the requirements for the degree
of Master of Arts in Psychology
at Massey University

Gail Robin Russell

1987

To

Louis, Sarah, and Kirsty



ACKNOWLEDGEMENTS

Very sincere thanks are due to my supervisors, Dr Judy Brook and Ms Shannon Roache for their guidance, support and encouragement, throughout the course of this research project.

I would also like to express my thanks to all the other people who contributed to the project: particularly Plunket Society staff and members; Joan Forrester who typed the questionnaires; Mavis Sievers who typed the final copy; and of course, my wonderful family who somehow managed to tolerate me.

ABSTRACT

The present study is an evaluation of the service delivery programme offered at the Palmerston North Plunket-Karitane family unit. The study had three aims: firstly, to replicate and extend an investigation conducted on a similar programme in Dunedin; secondly, to examine the aetiology and intensity of stress experienced by the service delivery staff; and thirdly, to systematically evaluate programme process and outcome. Results obtained in the present study were in many respects similar to those obtained in the Dunedin study, but some significant differences are also noted. Although valuable information pertaining to the causes of stress was obtained, the service delivery staff recorded stress levels comparable to other working women. Process and outcome evaluation data indicated that the programme was functioning in accordance with its aims and objectives, however recommendations for programme modification and improvement are offered.

TABLE OF CONTENTS

	Page
Acknowledgements	iii
Abstract	iv
Table of Contents	v
List of Tables	viii
List of Figures	xi
List of Appendices	xii
INTRODUCTION	
Brief History of the Royal New Zealand Plunket Society (Inc.)	1
The Development of Plunket-Karitane Family Units	2
Evaluation of the Service	3
CHAPTER ONE	
1.0 EVALUATION RESEARCH: AN OVERVIEW	5
1.1 The Problem of Definition	5
1.2 Research Design	7
1.3 Quantitative versus Qualitative Methods	8
1.4 Evaluation of Process and Outcome	8
1.5 Goal Setting and Evaluation	11
1.6 Goal Setting and Stress	12
1.7 Organizational Threats to Evaluation	14
1.8 Summary	15
CHAPTER TWO	
2.0 DEVELOPMENT OF CONCEPTUAL FRAMEWORKS	17
2.1 The Wortman Model	17
2.2 The Williamson, Prost and George Model	20
2.3 Relationship Between Models	23
2.4 Programme Integrity	25
2.5 Integration of Concepts	27
2.6 Summary	27

	Page
CHAPTER THREE	
3.0	METHODOLOGICAL ISSUES 29
3.1	In Evaluation 29
3.2	In Stress Research 30
3.3	Relevant Standardized Measures 32
3.4	Summary 33
CHAPTER FOUR	
4.0	RELEVANT STUDIES 34
4.1	Nurses and Stress 34
4.2	Part-time versus Full-time Employment 35
4.3	Social Support 35
4.4	Child Health in New Zealand 36
4.5	Summary 36
4.6	A New Zealand Study 36
CHAPTER FIVE	
5.0	THE PRESENT STUDY 42
5.1	Introduction 42
5.2	Replication Study 42
5.3	The Follow-up Study 43
5.4	Comparison Study 44
5.5	Stress Measure Development 44
5.6	The Stress Study 45
5.7	Needs Analysis 45
5.8	Summary 46
CHAPTER SIX	
6.0	METHOD 47
6.1	Replication Study 47
6.2	The Follow-up Study 48
6.3	Comparison Study 49
6.4	Stress Measure Development 50
6.5	The Stress Study 52
6.6	Needs Analysis 54
6.7	Other Sources of Information 55

	Page
CHAPTER SEVEN	
7.0 RESULTS	56
7.1 Replication Study	56
7.2 The Follow-up Study	68
7.3 Comparison Study	70
7.4 Family Unit Staff	71
7.5 Stress Measure Development	72
7.6 The Stress Study	73
7.7 Needs Analysis	78
CHAPTER EIGHT	
8.0 DISCUSSION	82
CHAPTER NINE	
9.0 CONCLUSIONS AND RECOMMENDATIONS	96
APPENDICES	100
REFERENCES	155

LIST OF TABLES

	Page
TABLE 2.01 Aspects of programme integrity	25
TABLE 7.01 Age of child on presentation at unit	56
TABLE 7.02a Socio-economic status of mothers presenting at family units in the present study and the Clarkson et al. (1985) study	57
TABLE 7.02b Socio-economic status of fathers in the present study and the Clarkson et al. (1985) study	58
TABLE 7.03 Size of families who presented at the family units in the present study and the Clarkson et al. (1985) study	58
TABLE 7.04 Source of referral to the family unit in the present study and the Clarkson et al. (1985) study	59
TABLE 7.05 Source of clients initial knowledge of unit in the present study	60
TABLE 7.06 Presenting problems in the present study and the Clarkson et al. (1985) study	61
TABLE 7.07a Frequency of usage of types of management (Clarkson et al., 1985 study)	62
TABLE 7.07b Frequency of usage of types of management (present study)	62
TABLE 7.08 Frequency of unit visits in the present study	64

	Page
TABLE 7.09	
Duration of contact with the Palmerston North family unit	64
TABLE 7.10	
Referral of family unit clients to other agencies in the present study	65
TABLE 7.11a	
Outcome evaluation - all problems combined (Clarkson et al., 1985)	65
TABLE 7.11b	
Outcome evaluation - all problems combined (present study)	66
TABLE 7.12	
Percentage of clients living at specified distances from the family unit	67
TABLE 7.13	
Presenting problems in the present study	68
TABLE 7.14	
Frequency of usage of types of problem management in the Follow-up Study	69
TABLE 7.15	
Age of survey respondents in the Comparison Study	70
TABLE 7.16	
Number of children of respondents in the Comparison study	70
TABLE 7.17	
Stressful incidents recorded in Nurses' Daily Stress Record	72
TABLE 7.18	
Stress measure data for experimental group and comparison groups for each data collection period	75
TABLE 7.19	
G.H.Q. 12 means and standard deviations for normative groups and present study groups	76

	Page
TABLE 7.20 Mean Daily Diary scores for experimental and comparison groups for each data collection period	77
TABLE 7.21 Experimental group responses to goal clarity probes in Goal Clarity Questionnaire	78
TABLE 7.22 Main aim of the Palmerston North family unit	79
TABLE 7.23 Subsidiary objectives of the Palmerston North family unit	80

LIST OF FIGURES

	Page
FIGURE 2.1	
The Wortman (1975) model of evaluation research	18
FIGURE 2.2	
Programme evaluation as a system feedback loop	20
FIGURE 2.3	
The relationship between the Wortman (1975) model, and the Williamson, Prost and George (1978) model	24
FIGURE 2.4	
The complementarity between the Williamson, Prost and George (1978), and Quay (1977) approaches to evaluation research	27
FIGURE 7.1	
Division of Palmerston North City into quadrants	67

LIST OF APPENDICES

	Page
APPENDIX A: Supplement to the Methods section	100
APPENDIX B: Supplement to the Results section. I.	145
APPENDIX C: Supplement to the Results section. II.	151

INTRODUCTION

Brief History of the Royal New Zealand Plunket Society (Inc.)

The Royal New Zealand Plunket Society (Incorporated) is a legally constituted organization whose members, also legally incorporated bodies, are the 119 Plunket Branch Societies throughout New Zealand. Each autonomous Branch embodies several sub-branches which functionally are sub-committees of the parent Branch, and are under its jurisdiction (Royal New Zealand Plunket Society, undated). The purpose of the Society is to provide a preventative health-care service to all New Zealand infants and pre-school children (Geddis and Silva, 1979).

The "Plunket Society" was founded in 1907 by Dr Fredrick Truby King, who was at that time Superintendent of Seacliffe Mental Hospital in Dunedin. Dr King saw a relationship between the high infant mortality rate; 73 per 1000 live births, (Parry, 1982) and the prevailing attitudes and practices associated with infant rearing. On the 14th May, 1907 he expounded his theories to an audience of influential Dunedin women, who enthusiastically embraced his doctrine of "... breast feeding, natural foods, cleanliness, and fresh air" (Travers, 1981, p.3). Within a year the Truby King Motherhood Movement, the forerunner to the Society for the Health of Women and Children, later to become the Royal New Zealand Plunket Society, was incorporated. The aim of the organization was to not only decrease infant mortality but also to improve the health of children through mutual support and education of parents. An early follower of the Truby King philosophy, Lady Plunket, wife of the contemporary Governor of New Zealand, gave her name to the specialist nurses employed to do this work, and later in recognition of her patronage the organization became known as the Plunket Society (ibid).

In the 80 years since inception the Plunket Society has become an integral part of the New Zealand health service. In 1980, Plunket nurses visited almost 42,000 new babies: 83% of all infants born in New Zealand that year (Parry, 1982), and a study of Dunedin families, (Geddis and Silva, 1979) reported 98% as having experienced some Plunket contact. The early years of the Society saw rapid expansion to provide a nationwide network of trained Plunket nurses who offered domicillary and clinic visits to parents needing practical assistance, advice, and support. By 1927 Karitane Hospitals had been established in Dunedin,

Christchurch, Auckland, Wanganui, Invercargill, and Wellington. The purpose of these hospitals was twofold: firstly, they served as training schools for Karitane nurses and Plunket nurses; and secondly, they were available for mothers and infants requiring more intensive residential care. Although the promotion of sound infant management practices has always been a primary concern, in the more recent past the Society has expanded its activities to include campaigning for hydatid eradication, bovine tuberculosis testing, fluoridation of water supplies, vaccination programmes, and child accident prevention.

Initially the Society was funded by donation, public subscription, and the fundraising activities of the volunteers who made up the Society's membership. As involvement in infant health care expanded, and running costs escalated, the Society was forced to seek Government assistance, however despite some pressure, it has retained its autonomy. Some funding is also provided by the Karitane Products Society Limited, a company set up in 1927 to manufacture the Plunket Society infant dietary supplements, which are only available for sale to Society members. Between 1979 and 1982 Karitane Products donated surplus profits of \$200,000 to the Plunket Society (Parry, 1982).

The Development of Plunket-Karitane Family Units

By the 1970s it became obvious that the role of Karitane Hospitals in the care of premature and weakling babies was being taken over by public hospitals with sophisticated neo-natal units. As Parry (1982) has noted "Karitane Hospitals belonged to a past era, when costs were lower, labour cheaper, mothers more ignorant, paediatric advice less readily available." (p. 158.) Furthermore, for many cases hospital care was inappropriate and Karitane Hospitals were underutilized and expensive to operate (Clarkson, Brown, Fraser, Herbison and Geddis, 1985). Parry (1982) records that in 1977 while 63% of the Plunket Society annual budget was spent on district work caring for 96% of the Plunket caseload, Karitane Hospitals consumed 37% to benefit just 4% of cases. Additionally, admission to the hospitals had become more commonly for social reasons, and less and less for ill-health, nutrition, or infant management problems. Between 1978 and 1980 Karitane Hospitals were phased out, and replaced by community-based family units, staffed by Plunket nurses and Karitanes.

The aim of the Plunket-Karitane family service is to provide an 'extended family' for parents "... offering helpful and well informed assistance with problems that often concern young parents with infants and toddlers. In practice, it offers a broad educational approach to child-rearing and family relationships." (Royal New Zealand Plunket Society information brochure.)

The Plunket-Karitane family service operates a dual system of assistance to families with young children through the family units and the mobile Karitane service. The family units provide an informal environment where parents can be advised on such matters as infant-care, household management, and family relationships. Units are also a place where those in need of a break from the demands of parenting can obtain a few hours rest, and those seeking social contact can meet other parents. Mobile Karitanes are available to go to family homes and give practical assistance with child-care and household management.

The Plunket-Karitane family service assistance is offered free of charge to its users, although voluntary contributions are accepted. It is funded partly by Government contribution, and the balance by public donation and funds raised by the voluntary sector of the Plunket Society.

Evaluation of the Service

Evaluation of this new type of Plunket service was obviously desirable once family units became well established. However, because clients in different areas have different needs, each of the 27 family units throughout New Zealand was given the autonomy to develop its own identity within the parameters of the family unit concept. This suggests that any indepth evaluation of family service functioning should be done at local level, since national evaluations yield only very general information. To date the only completed formal evaluation has been of the two Dunedin units (Clarkson et al., 1985).

The present study was conducted following an approach made to the Department of Psychology at Massey University by the Palmerston North Plunket-Karitane family unit management committee, for an evaluation study to be conducted. The management committee request was prompted by a resigning family unit nurse, who linked the high turnover of staff to the stress associated with working in a family unit.

Although a formal brief for the research project was requested, the management committee failed to provide one, leaving the extent and focus of the evaluation to the discretion of the research designers. The Palmerston North family unit was established in 1978. Because no formal evaluation of the unit programme had been undertaken since the unit had begun operation, a detailed evaluation research study, looking at programme resources, implementation, and effects was indicated. It had been suggested that stress was a factor contributing to staff resignations, and so while this area was targeted for particular attention, the evaluation was comprehensive and covered all aspects of the family unit functioning.

1.0 EVALUATION RESEARCH: AN OVERVIEW

1.1 The Problem of Definition

Evaluation as a legitimate field of applied social science research has become increasingly important over the last two decades (Raizen and Rossi, 1982): however, as Freeman (1977) and Riecken (1977) have both noted, opinion as to what constitutes evaluation research varies considerably. Riecken (1977) has observed that while some research which falls within the area of evaluation is not classified as such by its authors, other research which merely assembles statistical information, or provides informed judgement, is claimed to be evaluation. Glass and Ellett (1980) assert that definitions of evaluation, both good and bad, abound. They reproach writers for creating definitions to emphasize new aspects of evaluation, accusing them of being indifferent to whether a definition is too broad thereby including spurious concerns; or too narrow and excluding legitimate areas of investigation.

Glass and Ellett (1980) favour Scriven's (1967) definition, maintaining it captures the essential features of evaluation.

"[Evaluation] consists simply in the gathering and combining of performance data with a weighted set of critical scales to yield either comparative or numerical ratings, and in the justification of (a) the data-gathering instruments, (b) the weightings, and (c) the selection of criteria."

(Scriven, 1967, cited in Glass and Ellett, 1980, p. 212.)

Glass and Ellett (1980) consider most definitions of evaluation are stipulative, and applaud Scriven for a definition which approaches description, i.e. gives an account of prior usage. The present writer suspects the scope of Scriven's definition relegates it to the over-inclusive category suggested by Glass and Ellett (1980). Furthermore, this definition may well be incomprehensible to readers without prior knowledge of the dimensions of evaluation research.

A definition that limits evaluation research to assessment of human resource programmes through the use of social science methodology has been provided by Freeman (1977).

"... evaluation research is best defined as activities which follow the general mandates of social research, compromising these as minimally as possible because of the realities of the political and pragmatic environment in which investigators work. In other words, evaluation research is the application of social science methodologies to the assessment of human resource programs, so that it is possible to determine, empirically and with confidence that results from employing scientific procedures, whether or not they are useful." (Freeman, 1977, p. 25.)

Not only does this definition maintain reasonable boundaries for the field of evaluation research, but it also acknowledges the political and pragmatic constraints of applied research while stressing the need for empirical scientific methods. Unfortunately it falls short of being an acceptable definition by implying that the ultimate outcome of a programme is the only area of interest in evaluation research. (This issue will be discussed later in the chapter.) Freeman redresses the misconception in the Rossi and Freeman (1982) definition, used by Aiken and Kehrer (1985) to introduce a chapter addressing methodological issues in *Evaluation Studies Review Annual*.

"Evaluation research is the systematic application of social research procedures in assessing the conceptualization and design, implementation, and utility of social intervention programs... Evaluation research involves the use of social research methodologies to judge and to improve the planning, monitoring, effectiveness, and efficacy of health, education, welfare, and other human service programs." (Rossi and Freeman, 1982, p. 20.)

The present writer has adopted the Rossi and Freeman (1982) definition because explicitly or implicitly it embodies all the facets of evaluation research recognized by Scriven (1967) and Rossi (1977), while at the same time acknowledging that evaluation research incorporates assessment of process, as well as outcome. The definition stresses the need to examine programme rationale and design, programme operation, the overall usefulness of the intervention, and the use of information obtained to modify the programme toward the direction of the desired outcome. In short, the definition embraces all aspects of programme intervention, and not merely assessment of the end result.

The foregoing discussion of definitions serves to highlight the broad scope of evaluation research, and provides an indication of the many potential areas of debate within the field. A detailed 'state of the art' review of evaluation literature is beyond the scope of the present discussion: instead a discussion of selected issues of particular relevance to the present study is offered.

1.2 Research Design

The experiment has been defined by Keppel and Saufley (1980) as the basic means of establishing causal relationships between environmental events and specific behavioural effects. Cook and Campbell (1976) divide experimental design into two major categories: 'true' experiments, in which assignment to condition or treatment is random; and quasi-experiments, which approximate 'true' experiments in all respects except random assignment. Thus, in quasi-experiments subjects are assigned to conditions on the basis of already existing differences, (i.e. non-randomly) and therefore are non-equivalent before independent variable administration. Obviously, 'true' experiments permit the drawing of stronger causal inferences, but random assignment is not always possible, (or ethical) in applied research settings.

The constraints of applied research should not however, be used as an excuse for poor experimental control, and the field researcher needs to be equally concerned with proper research design as those who conduct laboratory experiments. Cook and Campbell (1976) offer suggestions for overcoming some randomization difficulties and note situations where opportunities for randomization might be maximized. However, in many cases the most appropriate design for a field project is quasi-experimental. Cook and Campbell (1976) discuss the four types of validity to be considered by researchers, and offer quasi-experimental designs incorporating controls to overcome some threats to validity. They also observe that while the ordering of forms of validity for the theoretician is internal, construct, conclusion, and external validity, for the applied researcher the priorities are ordered as follows: internal, external, conclusion, and construct validity. Wortman (1975) notes that for the applied researcher criteria for internal and external validity are the most difficult to satisfy. He has incorporated these dimensions, along with conclusion, and construct validity, in his model of evaluation research which is examined in Chapter two.

1.3 Quantitative versus Qualitative Methods

The relative merits of quantitative and qualitative methods has been the subject of protracted debate in the evaluation literature (Conner, Altman and Jackson, 1984). Because evaluation methodologies developed from the strategies used by social scientists in experimental research, early evaluation researchers were urged to view social reforms as experiments calling for quantitative methods (e.g. Campbell, 1969). Deshler (1984) credits Weiss (1972) with being one of the early supporters of qualitative methods in evaluation research, but as Conner et al. (1984) noted, the advocates of qualitative methods had to emphasize the differences between the approaches in order to be heard, and polarization of opinion resulted. Riecken and Boruch (1978) observed that while the proponents of the experimental approach stressed the need for internal validity and unbiased estimates of treatment efficacy, the critics pointed to limited external validity arising from the differences between experimental situations and the applied setting. However, they did report that progress was being made toward developing methods which utilize both qualitative and quantitative information.

Fortunately, as Conner et al. (1984) note, "The issue no longer seems to be which approach is better but, instead, is how we can capitalize on the complementarity of these approaches to design more sensitive studies." (p. 17.) These authors express a hope that rather than develop an amalgam of traditional approaches, the evolution of a unique evaluation research methodology will be encouraged. It seems important to reiterate here that the difficulties of evaluation research are no excuse for poor research design or methodology. The onus is on the evaluator to maximize the overall validity of research design.

1.4 Evaluation of Process and Outcome

Freeman (1977) reduces evaluation activities to two fundamental questions: firstly, whether the programme intervention was implemented as specified in the programme design (process evaluation); and secondly, whether the programme was effective in producing change (impact evaluation).

Conceptually it is probably easier to discuss evaluation of programme outcome (impact) first. Impact evaluation proceeds through analysis of the relationship between programme objectives and outcome variables, to

ascertain whether or not the intervention has achieved the desired and predicted results. Obviously, the nature of impact evaluation makes it readily adaptable to experimental or quasi-experimental design, using the treatment or intervention as the independent variable, and the outcome as the dependent variable. Freeman (1977) alleges that because experimental designs incorporate control of external biases, most evaluation researchers consider this to be the most appropriate way to measure programme impact. However, Christensen (1980) has warned that conceptualizing evaluation research as an experiment may imply "a degree of finality that it does not really have." (p. 287.) Elaborating on this statement he observed that impact evaluation includes the tacit obligation of the evaluator to furnish recommendations for programme improvement where necessary. Wortman (1975), following Scriven (1972), refers to the outcome phase of evaluation as summative evaluation, while Williamson, Prost and George (1978) have termed it goal-outcome congruence.

Outcome or impact evaluation assumes the programme intervention has taken place as originally prescribed, but this assumption is not always valid. Modification of the independent variable may occur for a myriad of reasons: those suggested by Freeman (1977) include unavailability of resources, political interference, poor motivation of staff, or failure to attract appropriate programme recipients. As Freeman (1977) noted, knowing the impact or outcome of a programme is of little value unless it can be shown that the programme did take place, and as intended. The assessment of how faithfully a programme has been implemented is referred to as process evaluation.

Two questions are central to process evaluation: first, has the programme reached the intended target population or target area? and second, were the various intervention procedures performed in compliance with programme design or derived from the ideals explicated in that design? (Freeman, 1977).

Process evaluation does not fit any particular paradigm, and is therefore more difficult to assess than impact evaluation. The evaluator is faced with the task of assembling data on programme operation and comparing current functioning with that specified in programme design. Similarly, description of programme recipients is compared with the significant variables of the target population. The assumption here is that

programme elements have been adequately defined and the target population accurately identified in the planning stages of the programme. If this was not done, the evaluator needs to have the programme designers clearly define the proposed programme inputs, including programme personnel, treatments, and programme recipients. Process evaluation includes Quay's (1977) concept of programme integrity: ensuring the treatment is of sufficient quality and quantity to meet programme requirements. Wortman (1975) used Scriven's (1972) term 'formative evaluation' to discuss programme inputs and processes.

Formative evaluation, which Wortman (1975) defines as "the process of treatment development or formation and the selection of goals" (p. 564), is a dynamic process. It not only looks at programme inputs, but also uses the data obtained from process-type evaluation as an information feedback loop to modify and improve programme operation. Christensen (1980) has likened formative evaluation to a series of impact evaluations, however the present writer considers a more accurate conceptualization would be as a series of process-impact evaluations, thereby giving proper emphasis to the evaluation of the formation of the programme. Williamson et al. (1978) have called this stage of evaluation means-end analysis.

A long-standing debate amongst evaluators centred around the relative merits of evaluations aimed at proving a programme's value and worth, versus evaluations aimed at improving a programme (Deshler, 1984). Deshler credits Cronbach with initiating the debate when in 1963 he argued that formative evaluation was more important and useful than summative evaluation. Freeman (1977) expressed similar views on the relative worth of process, and impact evaluations, while advocating that all evaluations should incorporate both process and impact components. Riecken (1977) too stressed the value of incorporating both formative and summative evaluation in research design, and Wortman (1975), and Williamson et al. (1978) have developed models which utilize both formative and summative evaluation strategies.

Early evaluation research was biased toward evaluation of programme outcome, and Freeman (1977) criticized evaluators for avoiding process evaluation because lack of technique refinement made it difficult to undertake. Freeman (1977) cited an investigation of federally funded studies (Bernstein and Freeman, 1975) which revealed that 25 percent

failed to measure process, and has suggested the percentage is higher in studies undertaken at local level both in the United States and abroad. Cronbach (1982) concedes that purely summative evaluation studies can be utilized where treatments are fixed, as in drug and vaccine testing. He remains sceptical about the value of pure summaries of outcome in social research, except where a programme is already fully developed. Brook (1984) notes that one type of outcome evaluation which has gained acceptance with the present emphasis on fiscal accountability, is that of cost-benefit or cost-effectiveness evaluation.

At the present time researchers appear to recognize the contribution to be made by both process and outcome evaluation. As Deshler (1984) states, "Today both kinds of evaluation are considered to be equally important, and to argue that one is better than the other is ridiculous." (p. 7.)

1.5 Goal Setting and Evaluation

Goals are a vital aspect of programme planning and evaluation. Wortman (1975) discusses formative evaluation as the development of the treatment intervention and selection of appropriate treatment goals; and summative evaluation as being concerned with how effectively the programme attains those objectives or goals. The selection of programme goals is rooted in the value system of a society, and a programme which operates successfully to achieve worthless goals is of little value. As Wortman (1975) eloquently stated "... qualitative intuition generates goals and precedes quantitative, statistical understanding of the progress made toward achieving those goals." (p. 565.)

In discussing goals, Zusman and Wurster (1975) assert that for some service delivery programmes, agency goals may be appropriate, but for others "... for example, an agency devoted to improving the quality of child rearing among its clients ..." (p. xviii) the primary objectives are not clear-cut, and the problem of goal setting complex. Kiresuk and Lund (1975) advocate examining goals at individual rather than organizational level: Goal Attainment Scaling was developed as a behavioural method of assessing progress toward operationally defined individual goals in clinical and therapeutic interventions, that could also be used to evaluate overall programme functioning. A critique of Goal Attainment

Scaling by Calsyn and Davidson (1978) commends the attempt to make evaluation more relevant to service providers, but concludes that Goal Attainment Scaling is differentially effective in meeting its dual aims. While conceding that there is evidence that Goal Attainment Scaling is effective as a therapeutic or programme management tool, they note that it lacks of psychometric properties, and should not be used as an evaluation technique in isolation. This procedure should be used as part of a multivariate strategy which includes valid measuring instruments administered to all clients in the programme.

1.6 Goal Setting and Stress

Goal setting has also been utilized in the management of stress. Stress was defined by McGrath (1976) as an imbalance between environmental demands and the organism's capacity to respond. He postulated six potential sources of stress in organizational settings: role-based stress, task-based stress, stress intrinsic to the behaviour setting, stress related to the physical environment, stress originating in the social milieu, and intra-personal stress. The Cooper and Marshall (1976) model also identified role-based stress as a major source of occupational stress.

In discussing aspects of role-based stress, Glowinkowski and Cooper (1986) identified three prime components; role conflict, role ambiguity, and having responsibility for others. Rizzo, House and Lirtzman (1970) defined role conflict as perception of inconsistent demands which may arise from:

- 1) inter-sender conflict when two or more persons make incompatible demands on the role holder
- 2) inter-role conflict occurring when a person holds two or more positions with incongruent demands
- 3) intra-sender conflict when time, resources, and personal capabilities are incompatible with the expected role behaviour
- 4) person-role conflict arising when personal values and defined role behaviour are incompatible.

Role ambiguity results when the role holder has insufficient knowledge of the role behaviour requirements of a given organizational position

(ibid). Glowinkowski and Cooper (1986) note that early researchers found a correlation between role stress, and physical and mental ill health. Later studies (e.g. Keenan and Newton, 1984; Martin, 1984) support this relationship, while Kemery, Bedeian, Mossholder and Touliatos (1985) found role ambiguity and role conflict directly linked to job satisfaction, job-related tension, and the tendency to resign in a group of accountants.

Newman and Beehr (1979), who conducted a major review of personal and organizational methods for dealing with job stress commented on a paucity of empirical evidence in the area, and the lack of contribution from industrial and organizational psychologists. Landy (1985) suggests four basic approaches to dealing with job stress: behavioural intervention, physiological intervention, cognitive intervention, and job design and redesign. Although empirical evaluation of the efficacy of the basic approaches in reducing occupational stress is scant, he cited two studies which support intervention at organizational level (Jackson 1983; Ganster, Mayes, Sime and Tharp, 1982).

Most of the literature on goal setting has investigated the relationship between goal specificity and performance (Quick, 1979). A review by Latham and Yukl (1975); and Quick's (1979) study suggest that goal setting activities are effective in reducing role ambiguity through goal clarification. Steer (1976) found a relationship between job satisfaction and goal specificity.

It has been suggested in the literature on stress management techniques that goal setting is a useful coping strategy, (e.g. Dewe, 1985; Muldary, 1983; Selye, 1974; Tubesing, 1981). Ganster et al. (1982) observed however, that while an intrapersonal approach was moderately successful in their study, changing environmental characteristics is more appropriate than teaching the individual to cope with stress induced by environmental demands. After reviewing the relevant literature, Murphy (1984) observed that most occupational stress management programmes are not aimed at reducing or eliminating sources of stress, but rather at teaching coping skills. He considers job redesign approaches to be preferable, but notes that there are significant problems associated with their development and implementation.

1.7 Organizational Threats to Evaluation

The basis of every social intervention programme is a value judgement, made by an individual or group, about the needs of some section of society. Wortman (1975) credited Glennan (1974) with the observation that social experimentation is a 'political act', recognizing that social programmes cannot be isolated from the political climate in which they exist. Evaluation research is primarily a 'political decision-making tool' (Christensen, 1980), making it incumbent on the researcher to not only design research from which valid conclusions may be drawn, but also to maximize the probability that research findings will be utilized.

Cook (1978) discusses evaluation research utilization at a theoretical level using Caplan's concept of two cultures. The first is a knowledge-generating culture seeking truth, validity, and goal-oriented rationality. The second is a potential knowledge-utilization culture valuing pragmatic action and process-oriented reality. The second culture stresses feasibility over idealism, timeliness over accuracy, and self-preservation over truth. Because the two cultures have different orientations, communication between them is difficult; the products of the former are seen as only marginally relevant to the perceived needs of the latter.

At the pragmatic level the researcher needs some awareness of the purpose of the evaluation. If those commissioning the research propose to use the information to justify a stance, findings contrary to their expectations are unlikely to be utilized. For example, Riecken (1977) observed that evaluation research may be invoked for a variety of purposes which include programme justification or improvement, but may also include attribution of blame for failure, leadership change, or curtailment of activities. He considers that serious research is not justified for the latter class of purposes since the intentions of those requiring the evaluation are preformed and unlikely to be influenced by research findings. Several authors (Agarwala-Rogers, 1977; Barnes, Brook, Hesketh and Johnson, 1985; Bonoma, 1977; Dowell and Kriesel, 1981; Williamson et al., 1978) offer practical suggestions to enhance the probability of research utilization. The recommendations include: providing feedback during the evaluation; presenting results in simple, clear language; acknowledging broader community constraints such as the

economic climate; and responding to criticism of the research openly and non-defensively.

Ethical considerations, always important in experimentation, are particularly relevant to social research where large numbers of participants and sensitive issues may be involved. Riecken (1977) suggests that the overriding principle should be that the intervention does not harm participants in any respect. A major responsibility of the researcher is to ensure confidentiality of information. Other ethical issues include the question of informed consent, and the withholding of treatment from the control groups. These issues are discussed by Riecken (1977), and Riecken and Boruch (1978).

The concept, of two cultures intercepting in evaluation research serves to highlight further difficulties encountered in evaluation studies. Programme staff may undervalue the evaluation effort, lacking motivation to obtain complete and accurate data, while at times programme values and needs must take precedence over evaluation needs. Staff may also resist evaluation efforts in the belief that they, and not the programme, are the primary focus of the research. Other problems include environmental constraints, difficulty in maintaining control over the experimental situation, and the problem of staff continuity. A further problem may be an increase in stress for staff involved in a study which could have job repercussions. Some of these difficulties may be overcome by working in a collaborative alliance with programme staff, but others may need to be accepted philosophically as concomitants of applied research. They should not however be used to justify inadequate research methods.

1.8 Summary

As the field of evaluation research has developed, the polarity of opinion that characterized discussion of many of the major issues has resolved. The novice evaluator, now left without strict "either/or" choices may however be left with the impression that in evaluation "anything goes". While it is accepted that each evaluation must be tailored to the individual organization, the guiding principle still remains sound research design tempered with imagination and common sense. Fortunately, some writers have drawn the requirements of sound evaluation research into models for research design, and while it is unlikely that

a specific model will meet every requirement of a particular project, these do provide an overall organizing conceptual framework from which to work. The use of particular models to develop a conceptual framework for the present study will be discussed in the next chapter.