NEWCASTLE UNIVERSITY LIBRARY
M 5 084 10892 3
Thesis L2901

THE EFFECTIVENESS OF BEHAVIOUR MODIFICATION PROCEDURES IN SECONDARY SCHOOLS WITH LIMITED TEACHER TRAINING AND CONSULTATION TIME

Submitted by Angus Macmillan to the University of Newcastle upon Tyne
as a thesis for the degree of Doctor of Philosophy
in the Faculty of Medicine.

AUGUST 1984

CONTENTS

PAGE

List of Tables..... 1

List of Figures..... 2

List of Appendices..... 3

Acknowledgements..... 4

Abstract..... 6

INTRODUCTION 7

(i) Why behaviour modification?..... 11

(ii) Why intervene within schools?..... 14

(iii) Behaviour modification within ordinary classroom settings..... 18

(iv) Intervening in the secondary school..... 23

(v) Conclusion..... 25

PART 1: THEORETICAL BACKGROUND 26

CHAPTER 1: MALADJUSTMENT: CONCEPTUAL AND ASSESSMENT ISSUES 27

(A) The Concept of Maladjustment 28

(i) What is maladjustment? Definitions and assumptions..... 28

(ii) The problems of maladjusted children..... 33

(iii) Prevalence..... 36

(iv) Persistence of children's problems..... 41

(v) Conclusion..... 43

(B) The Identification and Assessment of Maladjusted Children. 45

(1) Screening..... 46

(i) Issues of reliability and validity..... 46

(ii) Screening within schools..... 47

(iii) Single or multiple measures?..... 48

(iv) Ratings by teachers..... 49

(v) Ratings by peers - sociometry..... 53

(vi) Ratings by children..... 58

(vii) Conclusion..... 61

PAGE

(B) The Identification and Assessment of Maladjusted Children 62

(2) Observational Assessment

(i) Introduction..... 62
(ii) Methods of assessment..... 62
(iii) Units of observation and their definition..... 64
(iv) Observer training..... 66
(v) Computing inter-observer agreement..... 67
(vi) Establishing the validity of behavioural observation..... 68
(vii) Some sources of error in direct observation procedures... 69
(viii) Observer bias and expectancies..... 69
(ix) Reactivity of observational procedures..... 71
(x) Biases in the assessment of inter-observer agreement..... 72
(xi) Concluding note..... 73

CHAPTER 2: THE THEORETICAL BASIS FOR BEHAVIOURAL TREATMENT PROCEDURES 75

(A) The Philosophical Basis of the Operant Approach 77

(i) The emphasis on the study of individuals..... 77
(ii) The stance of radical behaviourism..... 81

(B) The Principles of Operant Behaviour 84

(i) General considerations..... 84
(ii) Procedures for strengthening behaviour..... 86
(iii) Procedures for weakening behaviour..... 90
(iv) Developing new and complex behaviour..... 91
(v) The role of antecedent events..... 93
(vi) Moving from the laboratory to less controlled settings... 96
(vii) Concluding note..... 98

CHAPTER 3: THE EVALUATION OF THERAPEUTIC OUTCOME 99

(A) Investigations of Therapy Effectiveness

(i) Child psychotherapy studies..... 101
(ii) Behaviour modification studies..... 108
(iii) Treatment comparison studies..... 111
(iv) Meta-analysis: an innovative approach to evaluating therapy outcome..... 114

PAGE

CHAPTER 3: THE EVALUATION OF THERAPEUTIC OUTCOME

(B) Methodological Issues 116

(i) Spontaneous remission..... 116
(ii) Experimental design..... 119
(iii) The assessment of change..... 123
(iv) The significance of change..... 126
(v) Assessing the maintenance of treatment..... 128
(vi) Conclusion..... 131

CHAPTER 4: REVIEW OF THE LITERATURE: TREATMENT INTERVENTIONS IN EDUCATIONAL SETTINGS 134

(A) Behaviour Modification Studies 135

(i) Interventions for inappropriate behaviour..... 137
(ii) Interventions for academic behaviours..... 172
(iii) Interventions for problems of social interaction..... 185
(iv) Changes in attitude and aspects of personality..... 191
(v) Issues of generalization..... 197
(vi) Training teachers in behaviour modification..... 206
(vii) The working hypotheses..... 219

(B) Consultation and Casework Studies 221

(i) School based consultation and casework..... 222
(ii) Casework studies..... 225
(iii) Conclusion..... 227

(C) Group Counselling Approaches with Children 227

(i) The nature of groups..... 228
(ii) How groups can be helpful..... 228
(iii) Group counselling with children's school-based problems..... 229
(iv) Conclusion..... 232

PART 2 : METHOD 233

CHAPTER 5: DESCRIPTION OF EXPERIMENTAL PROCEDURES 234

(i) Overview..... 235
(ii) Experimental design..... 236
(iii) The research hypotheses..... 238
(iv) The screen..... 240

PAGE

CHAPTER 5: DESCRIPTION OF EXPERIMENTAL PROCEDURES

(v) Allocation to regime..... 247
(vi) Organisation of regimes..... 248
(vii) Additional measures..... 253
(viii) Timing of assessments..... 256

CHAPTER 6: DESCRIPTION OF THE TREATMENT REGIMES 258

(A) Behaviour Modification 259
(i) The selection of teachers.....259
(ii) Training teachers in the application of behavioural procedures..... 261
(iii) Intervention procedures: general considerations.....264
(iv) Defining problems and establishing goals.....266
(v) The application of social reinforcement procedures.....270
(vi) Applying a behavioural contract system.....273
(vii) Provisions for maintenance of treatment effects.....274

(B) Parent Counselling/Teacher Consultation 275
(i) Training of social workers.....276
(ii) Teacher consultation.....276
(iii) Method of consultation.....278
(iv) Parent counselling.....279

(C) Group Counselling 280
(i) Rationale.....280
(ii) The structure of the regime.....281
(iii) The therapist training programme.....282
(iv) Therapy processes.....282

PART 3: DATA ANALYSIS AND DISCUSSION OF RESULTS 285

CHAPTER 7: RESULTS 286
(i) Methods of analysis (between-group).....287
(ii) Methods of analysis (within-group).....292
(iii) Outcome.....292

PAGE

CHAPTER 8: DISCUSSION 321

(i) Interpretation of findings..... 322

(ii) Problems implementing behavioural procedures in regular school settings..... 346

(iii) Critical views of behaviour modification..... 355

(iv) Methodological implications..... 368

(v) Practical recommendations..... 372

(vi) Summary..... 378

CHAPTER 9: CONCLUSION..... 379

APPENDICES..... 385

REFERENCES..... 427

LIST OF TABLES

TABLE	TITLE	PAGE
5 (1)	Basic design of research project.....	236
5 (2)	Cut-offs and weighting scores on screen criteria.....	245
5 (3)	Age and sex distribution of the children.....	248
5 (4)	Withdrawals from study before start of treatment....	251
6 (1)	Distribution of target children and teachers across schools.....	260
6 (2)	Distribution of target children and trained teachers across classes.....	261
7 (1)	Significant differences (paired comparisons): teacher ratings.....	294
7 (2)	Significant differences (paired comparisons): sociometric scores.....	296
7 (3)	Significant differences (paired comparisons): personality scores.....	297
7 (4)	Significant differences (paired comparisons): ability and reading comprehension scores.....	298
7 (5)	Significant differences (paired comparisons): Barker Lunn School Attitude Scale.....	300
7 (6)	Adjusted change means: teacher ratings (1).....	303
7 (7)	Adjusted change means: teacher ratings (2).....	304
7 (8)	Adjusted change means: sociometric, personality, ability and reading comprehension measures.....	305
7 (9)	Adjusted change means: attitude measures.....	306
7 (10)	Mean raw scores: teacher ratings (1).....	315
7 (11)	Mean raw scores: teacher ratings (2).....	316
7 (12)	Mean raw scores: sociometric and personality measures.....	317
7 (13)	Mean raw scores: ability and reading measures.....	318
7 (14)	Mean raw scores: attitude measures.....	319
A4(1)	Direct observation: pupil categories.....	415
A4(2)	Direct observation: teacher response categories.....	417
A4(3)	Observer agreement for task-relevant behaviour.....	418
A4(4)	Observer agreement for teacher approval and disapproval.....	420

LIST OF FIGURES

FIGURE	TITLE	PAGE
7 (1)	Adjusted change means: Rutter B2 scale.....	307
7 (2)	Adjusted change means: Devereux scale.....	308
7 (3)	Adjusted change means: isolation (sociometry) and neuroticism (JEPI).....	309
7 (4)	Adjusted change means: ability measures.....	310
7 (5)	Adjusted change means: reading comprehension..	311
7 (6)	Adjusted change means: Barker Lunn School Attitudes Scale.....	312
7 (7)	Adjusted change means: Barker Lunn School Attitudes Scale (continued).....	313
7 (8)	Levels of task relevant behaviour.....	314
A4(1)	Illustration of the rating sheet employed in direct observation.....	419

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
1	Screen measures.....	386
2	Establishing cut-offs and weightings for screen criteria.....	394
3	Second level measures.....	396
4	Observational procedures.....	413
5	Summary of content of training document and training seminars.....	421
6	Testing for homogeneity of within-group regression.....	423
7	Analysis of observational data on teacher behaviour.....	425

ACKNOWLEDGEMENTS

I would like at the outset to record my appreciation for the considerable help and support which I have received from my supervisor, Professor Issy Kolvin, throughout the conduct of the research and in the preparation of this thesis. His enthusiasm and commitment were invaluable qualities in sustaining the whole research programme, of which my project formed a part. My thanks are also due to the late Dr Roger Garside, deputy director of the research programme, for his advice, guidance and unfailing good humour. At the stage of writing up the thesis, I owe a particular debt to Dr Miller Mair at Crichton Royal, Dumfries, for his support, constructive advice, and the generous giving of his time. I am also grateful to my past and present colleagues Ray Ramsden, Peter Britton and Ron McKechnie for their helpful advice.

Thanks are due to the Department of Education and Science, who funded the research programme, and to the Dumfries and Galloway Health Board for a generous allocation of study leave to assist in writing up the thesis.

The co-operation and collaboration of many people are essential in a project such as the present one. Special thanks are due to the children who were the research subjects, and their parents; the teachers who helped in implementing the treatment programmes and in their evaluation; the head teachers and Directors of Education in Newcastle-upon-Tyne and Gateshead who allowed access to the schools. Thanks are due to the following people who were involved in different parts of the project: the directors of the treatment comparison regimes - Rory Nicol and Fred Wolstenholme; the fieldworkers in these regimes - Margaret Bell, Linda Harvey, Sue Jackson, Edith Khangure, Joan Parker and Liz Schlater; Linda Midgley-West and Nicola Whitehead for assistance with direct observations; Steve Iveson, Doris Fadden, Nanette Nelson and Isobel Smith for assistance with assessments; Linda Walker for help in developing the sociometric procedures; and Ian Leitch..../

Leitch for help with computer analyses of the data.

I would also like to express my thanks to Julie Oag for typing assistance and her tolerance and efficiency in coping with my handwriting through successive corrections and revisions.

Last, but not least, my thanks go to my family - Carol, Fiona and Karen - for their continued support, their acceptance of the time I have given over to studying and writing, and for their often-strained patience.

ABSTRACT

The aim of the study was to examine the effectiveness of behaviour modification procedures in alleviating the problems of maladjusted children in the first year of six secondary schools. Teachers (N=39) acted as mediators of treatment procedures, with a psychologist acting as consultant. Interventions were conducted within the constraints of regular school settings, so teacher training and consultation time were limited. Outcomes for the behaviour modification approach (N=74) were compared with those for untreated maladjusted controls (N=92) and two comparison treatment conditions: parent counselling/teacher consultation (N=83) and group counselling (N=73). Treated children and controls were selected by a multiple criterion screen and were randomly allocated by class to the various conditions. Multiple measures of change were employed, including classroom observations, teacher ratings, sociometry, and measures of personality, attitudes, verbal and non-verbal ability and reading comprehension. Follow-up assessments were conducted at three points in time, the last being three years after the initial screening. Significant differences in favour of behaviour modification were recorded on all change measures, mostly in comparison with maladjusted controls and parent counselling/teacher consultation, and such differences were observed at each of the follow-up points. Group counselling showed a similar pattern of outcome to behaviour modification.

In placing the present intervention in context, issues in the conceptualisation and assessment of maladjustment, and in the wider body of child therapy research are considered. The theoretical underpinnings of the behavioural approach are examined, and the relevant research literature in educational settings is reviewed. In discussion of the outcomes of the study, consideration is given to the complexity of the social and organisational context of such an intervention and to criticisms of the behavioural approach. On the basis of these considerations, recommendations for future interventions are offered.

INTRODUCTION

I N T R O D U C T I O N

Research into behaviour modification in educational settings has mushroomed in recent years. Programmes have encompassed a wide range of populations from preschoolers through to university students. A wide range of behaviours has also been altered, including performance on academic tasks, social behaviour and interaction both inside and outside the classroom, and a variety of specific problems which may hinder academic progress.

Although behaviour modification as an area of research inquiry is multifaceted in conceptual approaches and intervention procedures, applications within the schools have been based primarily on the use of rewarding and punishing consequences. Such procedures have been part of educational practices since time immemorial but the recent impetus for their use came from the development of operant conditioning and the work of B. F. Skinner. Given the long history of these procedures, the contribution of behaviour modification has perhaps been not so much in the development of new procedures as in their systematic application, refinement and evaluation within the classroom.

The present study was designed to investigate the effectiveness of behaviour modification procedures in helping maladjusted children, and to explore ways of conveying this help within ordinary schools rather than in specialised settings outside them. The nature and extent of 'maladjustment' will be examined in detail in Chapter 1. It has been estimated that between 7 and 25 per cent of school children/

children present with emotional or behavioural difficulties (Rutter et al., 1975; Kolvin et al., 1977; Macmillan et al., 1980). These problems may take various forms in the classroom or school, e.g. inattention, poor concentration and restlessness in class; poor academic motivation; disruptive behaviour; interpersonal difficulties with aggression and bullying on the one hand or shyness and inhibitions on the other, leading to isolation or rejection within the peer group. Not only are such children themselves often unhappy and miserable, but also much stress and strain is experienced by teachers who have to cope with them, often with little support (Kyriacou & Sutcliffe, 1978). Traditional approaches to childhood disturbance, such as individual psychotherapy , have not been conspicuously effective (Levitt, 1971) and behaviour modification offers one promising alternative. The present project attempted a controlled investigation of behavioural procedures within the regular school framework, which, it was hoped, would find substance behind the promise of behavioural research.

A major goal, then, was to convey and evaluate help offered within regular or 'natural' school settings - the issues involved here will be addressed shortly. But in addition, there are two fundamental questions which need to be asked, as they are basic to the orientation of the study. First, why was behaviour modification selected as the treatment of choice? Second, how may intervention within the schools be justified? Before attending to these questions, a brief account of the form of behavioural approaches in the classroom will help to put the subsequent discussion in context.

Behaviour modification programmes in the classroom have laid most/

most stress on the use of positive reinforcement procedures, where positive consequences are provided in order to increase certain desired behaviours. Positive consequences may take the form of social reinforcers (e.g. teacher praise, approval, attention, physical contact), activity reinforcers or privileges (e.g. preferred activities, games, free time) or token reinforcers (e.g. points, stars) which can be exchanged for other reinforcers such as activities, privileges or small toys. Punishment procedures have been employed to a much lesser extent, and they usually occur in conjunction with positive procedures rather than instead of them. They usually take the form of withdrawal of positive reinforcers. Thus, teacher attention or approval may be withheld, point losses may occur in a token reinforcement programme, or, in 'time-out' procedures, a child may, for a period of time, be denied access to available reinforcers, sometimes by removal from the situation.

These procedures are embedded in a model which stresses the clear definition of behaviours to be changed, the specification of objectives and how they are to be approached, and the re-structuring of environmental influences in order to bring about the desired modification of behaviour. In most applications, a 'triadic' model is involved, with the teacher as the 'mediator' of procedures, the psychologist as a 'consultant' and the child as the 'target'. It is evident that such a model often attempts to create change in teacher behaviour before change can be expected in the child's behaviour. More detailed consideration of the nature of classroom behaviour modification and its effectiveness will be given in Chapter 4. Comprehensive reviews are available in Sherman & Bushell (1974), O'Leary & O'Leary (1976), Nietzel et al., (1977) and/

and O'Leary (1978).

(i) Why Behaviour Modification?

As an approach to the problems of maladjusted children in schools, behaviour modification, for a number of reasons, seems an appropriate choice of treatment modality. A fundamental reason is that it can be viewed as generally compatible with many teaching activities. Not only are applications of behavioural procedures practicable on a group or class-wide as well as an individual basis (Litow & Pumroy, 1975), they can be viewed as having an affinity with basic educational principles. Many of the teacher's concerns, such as carefully observing pupils' behaviour, specifying objectives for learning, arranging for these to be gradually and systematically approached, and harnessing pupils' motivation, are also those of the behaviour modifier. For the teacher confronted with classroom management problems, behaviour modification offers a wide range of possible intervention strategies. This is not to claim that it is a panacea for all ills, nor that it is simply a cookbook collection of techniques. Rather it can be viewed as offering a conceptual framework with which to approach problem situations and within which a number of treatment options are available. It is also, of course, just one possible conceptual framework, and as such, it may not prove to be consistent or compatible with the beliefs or ideologies of individual teachers.

One of the most compelling reasons for recommending behaviour modification as an approach worthy of exploration is its primarily positive orientation. There are several ways in which it can be regarded as 'positive' . First, in the context of discipline problems,/

problems, it seeks to shift the balance away from the common reliance on verbal disapproval, reprimands and occasional corporal punishment in favour of methods which stress positive consequences for desirable behaviour. By so doing, it may help to make classroom management more humane and less coercive. There is little doubt that such a move would improve classroom climate. Patterson et al., (1969) have characterised the American classroom as a 'barren wasteland when one compares it with other normally reinforcing interactions... The control of social behaviour is achieved more as a function of threatened or applied aversive consequences than by positive social reinforcers'. This impression is backed up by data such as those presented by White (1975), who found that, after the first and second grades, the use of approval diminishes and is consistently exceeded by the use of disapproval. Similarly, a large survey of public school teachers in Florida showed that 77 per cent of their interactions with children were negative (Madsen et al., 1970). Both children and teachers would no doubt welcome ways in which their interactions could become more positively toned.

A second positive feature of behaviour modification is its 'constructional' approach to problems. This is defined as 'an orientation whose solution to problems is the construction of repertoires (or their reinstatement or transfer to new situations) rather than the elimination of repertoires' (Goldiamond, 1974, p.14). Rather than dwelling on a child's deficiencies or weaknesses, a constructional approach focuses on what repertoires, strengths and skills are available, and uses these as bases for taking the child in the desired direction. It may be contrasted with an 'eliminative' approach/

approach in which presenting complaints are classified as behavioural excesses and intervention procedures are designed solely to get rid of the problems without attempting in any way to replace them with more appropriate behaviours. In simple terms, eliminative programmes show children what not to do, while constructional programmes show them what they could be doing instead.

A third way in which a 'positive' orientation is evident is the manner in which a reconceptualisation of teaching failures is encouraged (Macmillan, 1973). When children fail at a learning task, 'explanations' are frequently couched in negative terms which imply some kind of deficit or inadequacy in the child. The behavioural approach, with an emphasis on task analysis, sequencing activities in small increments and giving effective reinforcers contingent upon successful performance, forces an alternative conception of failures. By heavily stressing the role of environmental arrangements, an absence of learning can be seen as being as much a teaching failure as a child failure.

Taken together, these three aspects of behaviour modification suggest an encouraging, positive re-orientation towards children's problems in school, and given the way in which the approach complements regular teaching activities, it seems well suited to classroom applications.

One implication of the foregoing discussion is that a shift of responsibility - or 'blame' - for children's problems is involved, with the onus being placed more on the teacher than the child. Certainly many of the preceding points imply that if children's behaviour/

behaviour or performance is going to change, some aspect of the teacher's behaviour has to change first. It may be that the initiative in dealing with problems rests, in most cases, with teachers but many of the problems which arise in classrooms are problems of relationships or interactions (Gillham, 1981). They occur between people rather than within them and their resolution requires an interactional framework, not one which simply stresses one-way paths of influence. Behaviour modification would seem compatible with such a perspective in that it stresses reciprocal influences, with each party to a transaction being seen as influencing, and reacting to the influences of, the other party. The effectiveness that is claimed for these procedures, is, of course, an especially attractive feature, and the evidence for such claims will be examined in Chapter 4 . Behaviour modification, therefore, may well be an appropriate choice of treatment but now the more fundamental question will be addressed - why should such an intervention be located within the schools?

(ii) Why Intervene Within Schools?

There has been an increasing recognition in recent years that schools need to carry more of the responsibility for influencing and changing pupils' behaviour - responsibility which in the past tended to be left to outside agencies. As a result, more is expected of teachers, and of those involved in the organisation of schooling. The thrust of the present study was very much in line with this shift in thinking. Where children with special needs are concerned - including the maladjusted - these changing expectations were enshrined recently in this country in the Warnock Report (D.E.S., 1978) which called on ordinary schools to share more of the tasks of/

of 'special' education. Schools inevitably have socialising functions as well as the more obvious academic and curricular ones, and clearly have an impact on the emotional and behavioural aspects of children's lives. Schools can be viewed as a microcosm of society outside (Duke, 1979) and as transmitting prevailing attitudes and values. This may be seen as an implicit process, which few would question, but there is heated debate on the issue of whether the school's socialising and helping potential should be vigorously acted upon and realised. Thus, while some teachers would welcome greater emphasis on pastoral activities to meet the needs of maladjusted children, others see solutions, especially where disruptive children are concerned, as lying in special provision outside of the ordinary classroom or school framework.

Apart from this shift in educational policy towards widening the role of ordinary schools, there are other thrusts in the same direction which without question helped to inform policy changes. First, following the rather demoralising studies emanating from the United States between the mid-60s and the mid-70s suggesting that schools had little impact on student attainment and performance as against the overriding influence of the home and the community (e.g. Coleman, 1966; Jencks 1972) evidence to the contrary emerged in this country. Thus Power et al., (1972) and later, Reynolds (1976) and Gath et al., (1977) presented findings which purported to show wide differences in delinquency rates (amongst other factors) between schools which drew from similar catchment areas. Reynolds concluded/

concluded from his study that 'these nine schools therefore appear to be producing large differences in the rates of delinquency, truancy and academic attainment of their pupils, differences which do not appear to be significantly related to variations in the social background of the catchment areas from which the schools take their pupils'. More recently, Rutter et al., (1979) in a study of twelve London comprehensive schools similarly found consistent differences between schools on delinquency rates, attendance, within-school behaviour and public examination results, and these differences remained after differences between the schools in the quality of their intake were taken into account. Although this study has been vigorously criticised (e.g. Goldstein, 1980; Hargreaves, 1980) few would dispute the authors' conclusion that 'the results carry the strong implication that schools can do much to foster good behaviour and attainment and that even in a disadvantaged area schools can be a force for the good.'

In addition to this confirmation of schools' potential influence and impact on children - for good or ill - changing conceptions of the nature of childhood disturbance or mal-adjustment and awareness of the inadequacies of traditional treatments, began to stimulate new intervention models. These models tended to reinforce the notion of schools as one community setting within which intervention efforts could be appropriately located. The changing view of childhood disturbance represented a shift from conceptions of problems as intra-psychic, child-centred and non-situationally specific, to ones in/

in which disturbed behaviour was increasingly seen as a function of interpersonal and social contexts. Where the traditional perspective encouraged a belief that individually-oriented treatment by outside specialists was the intervention of choice and alienated teachers from active and constructive involvement (Leach, 1977) the emerging, essentially interactive perspective invited school based personnel to take a significant role in assisting children's adjustment. These moves were encouraged by the apparent failures of traditional helping methods. Thus, workers reviewing individual psychotherapy concluded that it was ineffective, or, at best, that its case was not proven (Eisenberg, 1969; Levitt, 1971). A similar crisis of effectiveness was evident in the Child Guidance Service, the main provision for helping troubled children in the schools (Gillham, 1981). Shepherd et al's (1971) study of school children in Buckinghamshire suggested that, over a three-year period, those who received child guidance treatment showed identical rates of improvement to those who did not.

In combination, these different 'thrusts' can be seen as compelling greater attention to the contribution that schools themselves can make to solving the problems of their charges. There are indeed many advantages to the widening of teachers' roles, and giving explicit recognition to functions other than the mere purveying of academic knowledge. First, teachers are well placed to convey help, in that children spend about half their waking day in contact with them and about 15,000 hours/

hours in the course of their school career, according to Rutter et al. (1979). Second, they are exposed to large numbers of children, and if, as it is frequently complained, mental health professionals do not exist in adequate numbers to cope with those in need, intervention methods employing teachers as mediators are attractive. Third, if children can be helped within the ordinary school, rather than at clinics or day or residential special schools, stigmatising labels are more likely to be avoided. There is, in addition, little evidence that such provision outside ordinary schools is effective (Galloway, 1979). Finally, if teachers' helping and management skills can be improved, this may have important preventive implications so that not only may current problems be dealt with but also developing ones may be forestalled.

School-based interventions form part of the more general shift in recent years towards community-based programmes in which the helping potential of teachers, parents and others has been exploited. From a research perspective, rigorous evaluation of such interventions is as necessary as of their more traditional counterparts. These new settings also create special difficulties and hazards for the researcher who finds himself on unfamiliar ground. Some of the questions posed about the transition into 'natural' environments provided the stimulus for the present project, and these will now be examined.

(iii) Behaviour Modification Within Ordinary Classroom Settings

Despite the success that has been claimed for behavioural procedures/

procedures, a number of questions and critical issues can be raised, and some of these are as relevant now as they were at the time the present study was being planned. A fundamental concern, to which the present study was addressed, was the extent to which successful outcomes might be dependent on the research context in which the work was carried out, which could be viewed as limiting their relevance to ordinary school settings. There are a number of facets to this issue which helped to shape the present investigation, and these will be outlined here.

Several features of research settings may influence the findings of studies, yet tend often not to be acknowledged as important variables. Thus it has been suggested by O'Leary & Kent (1973) that features such as course credits or joint authorships for teachers in return for their participation, and frequent monitoring of progress by research personnel constitute advantages unavailable in ordinary school settings but which may contribute substantially to the positive outcomes reported. In addition, ready availability of research funds gives more flexibility in choice of reinforcers for children, with the opportunity of maximising motivation. Considering motivation at teacher level, commitment to get change may be considerably increased where the study constitutes work for a higher degree. These features can readily be seen as differentiating the research context from the context in which educational and clinical psychologists function in relation to/

to teachers and problem children.

An important aspect of the published literature, up to the time the present investigation was being organised (1973-74) was the extent to which studies were designed as 'demonstration' models. In other words, the concern was to show that behaviour modification or a specific procedure within the behavioural umbrella actually 'worked'. This desired outcome may be achieved most easily not by intervention models which closely replicate regular or 'natural environment' conditions, but by ones which, by incorporating factors such as the above, to a considerable extent depart from such conditions. A major concern was that such investigations, intent on displaying the impact of one or other specific procedure, may focus on behaviours which are readily controlled and changed but which may be of limited practical significance. It was certainly apparent that early behavioural studies were preoccupied with measures of observable behaviour within the classroom to the virtual exclusion of other measurement modalities, such as those dealing with cognitive and attitudinal changes, or with the perceptions of significant parties such as teachers and peers. A further possible drawback of 'demonstration' models is that they may dwell on short-term changes at the expense of investigating change and maintenance in the longer term. Where long term follow-up is concerned, the field as a whole suffers from a dearth of data (Keeley et al., 1976).

For practitioners working in service settings, these issues could be seen as important limitations. In an attempt to meet them, the present study was designed to include multiple measures of change, including observational measures, and to cater for long-term follow-up/

up. It is of interest to note that the concerns addressed here have been persistent ones, as is evident from recent commentaries on classroom behaviour modification as a whole (e.g. Kazdin, 1981a)

We have considered some of the hidden, but possibly influential features of the research environment - what are the important 'natural environment' conditions? In this country, at least, they tend to be characterised at the general level by inadequacy of professional manpower relative to the numbers of children considered to be in need of help (Kolvin et al., 1973). Apart from implying pressures on mental health workers, these facts direct attention to the need for innovative methods of intervention which can reach larger numbers of those requiring help. Such interventions may bring their own special problems, given their location in the complex social setting of the school. Due consideration, for example, must be given to the manner in which a programme is introduced to the schools, if its entry is to be smooth, and also to its subsequent support and maintenance.

Psychologists wishing to undertake behavioural interventions may be hampered by a number of problems which their research counterparts successfully elude. Thus, they may have to employ as mediators teachers who are less than favourably disposed to behavioural methods and who may not be well motivated to carry them through. They may have to forego the luxury of independent observers who can provide the sort of data that is the cornerstone of the behavioural approach. Where there are no funds or special resources channelled towards the intervention, the impetus that attaches to many authorised research projects is missing.

A/

A particular difficulty for workers in regular settings, and one which bears on the adequate implementation of treatment procedures, is that sufficient time for teacher training and consultation during the course of the programme, may not be available. One of the particular tasks of the present study was to evolve a 'training' and consultation model which could operate within the time constraints of the ordinary school without prejudicing the quality of the treatment interventions. At the time of planning, reports of the demands on professional time for teacher training were quite daunting. Thus, Abidin (1971) reported that school psychologists could expect to spend approximately 30 hours assisting teachers with no previous experience in behavioural techniques to institute an individual programme and 150 hours to establish a token economy. Similarly, Patterson et al (1973) reporting on a series of successfully completed cases, indicated that, on average, 27.5 hours of professional time was necessary. Established training courses such as the Consulting Teacher Model (McKenzie et al., 1970) and the Responsive Teacher Programme (Hall & Copeland, 1972) were constructed around a degree or extended in-service training format entailing considerable demands on organisation, time and resources. Some studies were available which reported successful outcome with limited training and consultation time (e.g. Andrews, 1970; Kuypers et al., 1968) but these were restricted to small numbers of teachers and target children. In brief, there were no ready-made models available which could be translated directly to large-scale interventions in regular settings. The issues involved in teacher training will be examined more fully in Chapter 4.

(iv)/...

(iv) Intervening in the Secondary School.

In addition to the reservations noted above which raise questions about the general applicability of the behavioural approach, it was also apparent that the secondary school population had been neglected, with interventions having been undertaken almost exclusively with primary school children. There are a number of important differences between primary and secondary schools which suggest that the reported outcomes in the former settings may not necessarily hold for the latter. First, there are important differences in teacher-pupil contact. In the primary school, the one teacher will spend most of each day with the one class. His secondary school counterpart may take a class once a week, or perhaps five or six times a week. In the course of a week, he might teach 200-300 children across different classes, so he may have limited opportunity to get to know individual children well. From the pupil's point of view, he might be exposed to up to fifteen different teachers in a week. Consistency of management across staff, an important consideration for behavioural programmes, is therefore a potential difficulty in secondary schools - the issue may easily be avoided in primary schools. Where only a few secondary school teachers can be enlisted, as part of a behavioural programme, is this adequate? Will their involvement make enough of an impact on the pupil who is confronted with an array of teachers who behave towards him in different ways?

A second difference between primary and secondary settings is that the greater tendency in secondary schools for teaching to assume a 'lecture' format, with the teacher talking from the front of the class, may not be conducive to frequent teacher praise, attention/..

attention or physical proximity which may be required, say, in a social reinforcement programme. Can these procedures therefore be applied readily in the secondary classroom? A third, and related, difficulty is that even if such procedures can be applied, are they as effective with secondary-age children as they are claimed to be with younger children? It has been suggested (e.g. McNamara and Harrop, 1979) that teacher praise and attention may not be of great value with secondary aged children because this may for some pupils alienate them from the peer group. Inviting the teacher's disapproval, or 'bucking the system', may be more attractive within the peer group context than any reinforcer the teacher can offer. Much of this is speculative, since, as will be evident from the literature review, to be presented later, very few studies of social reinforcement in secondary schools have in fact been reported. This was one little-researched area in which the present study sought to make a contribution, with social reinforcement being given the major emphasis.

A fourth difference between the two settings is that secondary schools are perhaps more rigid and inflexible in terms of timetabling and organisation, and reinforcers other than teacher attention and praise (e.g. preferred activities, free time) are less easy to mobilise and make available than in primary schools.

These issues raise a number of questions, therefore, about the viability of behaviour mofidication in secondary school settings. In being concerned with children in such schools, it was hoped that the present study might help to provide some answers to these questions.

(v)/...

(v) Conclusion

In conclusion, behaviour modification was felt to be an appropriate model of treatment to apply within ordinary school settings, but the choice of such a location for helping efforts creates a variety of practical and theoretical problems. It was the task of the present study to attempt to meet these difficulties, and to evaluate the effectiveness of the help being conveyed. The study had a number of features which were seen as potentially important contributions, by tackling gaps in the research literature or areas which were relatively neglected. These features may be summarised as follows: (a) an attempt to conduct treatment interventions within ordinary school settings, with a minimum of research 'trappings', (b) examining the use of behaviour modification in the relatively unexplored context of the secondary school, (c) examining the impact of systematic social reinforcement with secondary aged children - again, a little-researched procedure in this setting, (d) employing multiple measures of change, so that functioning in a number of domains could be investigated and (e) conducting long-term follow-up to examine the impact of intervention beyond the treatment phase and its immediate aftermath.

Before these features, and the design and organisation of the study can be described in greater detail, the next four chapters will provide the necessary theoretical background. These will examine the nature and assessment of 'maladjustment' and the theoretical underpinnings of behaviour modification, consider the wider context of child treatment research, and, finally, review the relevant literature in educational settings.

PART 1

THEORETICAL BACKGROUND

CHAPTER 1

MALADJUSTMENT: CONCEPTUAL AND ASSESSMENT ISSUES

MALADJUSTMENT: CONCEPTUAL AND ASSESSMENT ISSUES

The present study was primarily concerned with conveying help to maladjusted children within school settings. This chapter will provide a context for understanding how these children's problems may be conceptualized, by attempting to clarify the nature of 'maladjustment' and the kinds of problem being dealt with. Some justification for the use of the term will be offered, since it has been subjected to criticism. Emphasis will be placed on consideration of how maladjusted children might be identified since an important part of the present project involved the identification as well as the treatment of such children. In addition, consideration will be given to issues in the area of direct observational assessment of behaviour, since measures of this type were integral to the treatment intervention, and were used, amongst other information, to complement the information gathered by the initial identification (screening) procedures.

The first section will examine: issues to do with the concept of maladjustment and its associated assumptions; the kinds of behaviour problems maladjusted children show; the prevalence and persistence of such problems.

(A) The Concept of Maladjustment

(i) What is maladjustment? Definitions and assumptions.

One of the most striking aspects of the writings on troubled children and their difficulties is the plethora of terms and labels used to describe them. Their problems are variously referred to as "maladjustment", "emotional disorder"/"disturbance"/"handicap", "conduct" or "antisocial disorder","psychiatric disorder" and so on. The term used/

used throughout this study to refer to the troubled children who are the focus of concern is "maladjusted". The term has been in common usage since the Second World War when maladjusted pupils were defined in the Handicapped Pupils and School Health Service Regulations, 1945, as "pupils who show evidence of emotional instability or psychological disturbance and require special educational treatment in order to effect their personal, social or educational re-adjustment". The Underwood Committee (Ministry of Education, 1955) accepted this legal definition of maladjustment, seeing it as making possible special education provision for a wide range of children with emotional or behavioural difficulties. However, its failure to offer any help in the identification of maladjustment in particular children was noted and it is evident that the substitutive nature of the definition offered - "emotional instability" and "psychological disturbance" themselves requiring clarification - leads to circular reasoning. The Underwood Report went on to emphasise that maladjustment described a relationship between the individual and his environment and that the maladjusted child was "developing in ways that have a bad effect on himself or his fellows and cannot without help be remedied by his parents, teachers and the other adults in ordinary contact with him." Lovell (1958) took issue with this, in stressing that there are degrees of maladjustment and that we should also be concerned to help those children whose maladjustment is slight. The Underwood Committee also adopted a medical model in viewing children's problems as "symptoms" and Galloway and Goodwin (1979)/

(1979) have pointed to the illogicality that special educational treatment is needed for children with quasi-medical symptoms of emotional instability and psychological disturbance. Working from an explicitly psychiatric/ medical base, Rutter and Graham (1968) offer a definition of psychiatric disorder which is frequently cited in the field and which can be accepted as a workable definition of maladjustment. In their terms:

"Psychiatric disorder....refers to abnormalities of emotions, behaviour or relationships which are developmentally inappropriate and of sufficient duration and severity to cause persistent suffering or handicap to the child and/or distress or disturbance to the family or community."

They go on to suggest that their use of the term "psychiatric disorder" does not involve any concept of disease or illness, nor does it necessarily assume that psychiatrists are the right people to treat such disorders.

The term "maladjusted" has been retained in the present study partly because of its familiarity in the literature, despite its difficulties, and it is intended only as an umbrella term, requiring more precise and individual specification for particular children. More importantly, it is retained because of the important implications it has, in its correct usage, in compelling us to attend to what the child is maladjusted to. Without a consideration of this essential relational component, the term is less than helpful, and it must be acknowledged that the term may not be used by some in this discriminating fashion. There is a sense in which most definitions of maladjustment, and perhaps most obviously those which refer to problems as "disturbing others" have a rather confused status. On the one hand they are suggestive of objective/

objective empirical properties, while on the other they imply judgments that are essentially value-appraisals. The extent to which the latter influence our thinking tends often not to be made explicit or even acknowledged. Labelling theory, within sociology (e.g. Hargreaves, 1978), suggests that the kind of deviance that may be implied by the term 'maladjustment' is very much a social construction. Erikson (1962), for example, writes:

"Deviance is not a property inherent in certain forms of behaviour: it is a property conferred upon these forms by audiences which directly witness them. The critical variable in the study of deviance, then, is the social audience rather than the individual actor, since it is the audience which eventually determines whether or not any episode of behaviour or class of episodes is labelled deviant.' (p.309)

In its strongest form, labelling theory suggests that labelling a child as deviant - or, presumably, 'maladjusted' - creates or amplifies deviance (Hargreaves, 1978). Without necessarily subscribing to this position, it is important to be aware of the extent to which differing perspectives and value judgments may help to determine what is identified as 'maladjustment'. As will be shown in a later section, there are many problems at the level of definitions and operations in the assessment of maladjustment, and differing conventions contribute to varying estimates of prevalence.

There are a number of assumptions which often flow from the use of terms like 'maladjusted' and other such labels for problem or disturbed behaviour. These need some clarification, since some of them are not consistent with the writer's stance towards intervention in this study. These discordant assumptions are ones associated with a 'medical/

'medical model' of understanding. Within this perspective, maladjustment or abnormal behaviour is viewed as analogous to physical disease: just as physical symptoms often result from some underlying physical pathology, so abnormal behaviour is thought to often be symptomatic of underlying pathology (Schwartz & Johnson, 1981). The treatment focus will be on resolving the underlying causes which might take the form of fixations, unresolved intrapsychic conflicts, personality traits and so on. Begelman (1976) additionally points out that this model would view pathology as residing within the individual rather than in the external environment or social structures influencing his behaviour, and that these internal processes enable one to predict behaviour across different situations. Psychodynamically oriented approaches to therapy tend to be based on this model, which until recently provided the most widely accepted framework for understanding deviant behaviour.

An alternative conceptualization of abnormal behaviour, with which the treatment in the present study was identified, is the behavioural approach. In this model, deviant behaviour is viewed as understandable in terms of learning history and environmental factors which serve to elicit and maintain it. The contribution of biological factors is not, however, ignored. Where treatment is concerned, emphasis is placed on external environmental influences on maladaptive behaviour rather than on intrapsychic ones. Situational determinants of behaviour are stressed, in contrast to the stable dispositions or traits of the medical model, so changes in one setting are not viewed as necessarily having implications for changes in others. It is inappropriate therefore, on this view, to assume that 'maladjusted' children would display their maladjustment ubiquitously.

This is a brief summary of the basic assumptions of these two contrasting models. As Schwartz & Johnston (1981) point out, there is in actuality no such thing as 'the medical model' or 'the behavioural model'. These sets of assumptions should merely be seen as a guiding framework within which adherents can vary widely. Thus, behavioural conceptualizations can range from those who focus exclusively on overt behaviours to those who emphasise cognitions as mediators of behaviour. However, the principal objective in introducing these two models here is to make clear the dissociation between the conceptualizations underlying the behavioural approach of the present study, and those associated with the medical model. This is done in the belief that it is possible to employ terms like 'maladjustment' and 'maladjusted child' without incorporating medical-model assumptions. As Chazan (1963) suggests, 'provided it is used not as a medical entity but as an umbrella term which is administratively convenient, there seems little reason for changing a term which now covers such a wide variety of psychological conditions'. The relational component, referred to earlier, in which the child is regarded as 'maladjusted to' in fact implies some referent in the child's environment, which is very much in keeping with the behavioural model advanced here.

(ii) The problems of maladjusted children.

What kinds of problems are we concerned with in dealing with maladjusted children? What is their prevalence? How long do they last? These are important questions to consider in contemplating a treatment intervention. The answers should assist decisions as to whether treatment is necessary, with what intensity and in what form.

Traditional/

Traditional approaches to classifying children's problem behaviour have been based on clinical observation, in which diagnoses are arrived at on the basis of particular constellations of symptoms. More recently, multivariate statistical procedures have been employed (Quay, 1979). Typically, factor analytic techniques are employed to determine clusters of behavioural characteristics that are highly correlated (e.g. on behaviour checklists) and seem to be commonly found together. Assuming an adequate sampling of behavioural characteristics and an adequate sample of the population, this approach is assumed to allow the determination of the major patterns of maladaptive behaviour.

As pointed out by Herbert (1978), there is an impressive consensus among studies of childhood problem behaviours using multivariate methods about the reality of two constellations of problems. On the one hand there is the cluster which includes physical and verbal aggressiveness, disruptive and destructive behaviour, non-compliance, stealing, and poor interpersonal relationships. Ths pattern is usually referred to as 'conduct' or 'antisocial' problems. On the other hand, there is a cluster which includes feelings of anxiety or depression, shyness, somatic complaints, day-dreaming, crying and special fears. These are usually referred to as 'neurotic' or 'anxiety-withdrawal' problems. The distinction between conduct and neurotic problems can be seen as referring to the direction of the maladaptive response: thus, Peterson (1961) suggests that:

"In one case, impulses are expressed and society suffers; in the other case, impulses are evidently inhibited and the child suffers'.

These two clusters have appeared in several empirical/

empirical studies carried out with children in different countries, in general population, child guidance or residential samples, and with younger or older children (Achenbach, 1966; Robins 1972; Kolvin et al., 1975). While most children with difficulties present with either a predominantly neurotic or conduct clustering, there are several other presentations with which we are not concerned here (see Rutter, 1965). These include, for example, psychotic states, developmental disorders and hyperkinetic conditions.

From an examination of the contents of the above two main clusters, it will be evident that many of these behaviours will be manifest within the school and within the classroom. This is an important consideration for intervention strategies which are designed to be school based - as in the present project. Such interventions clearly have to 'reach' their intended target. However, problems of disobedience, aggression and disruptive behaviour, and their associated difficulties of poor concentration and attention span, are no rare occurrences for teachers. The problems of the anxious, fearful and shy child may not be as obtrusive as those of his acting-out counterpart, but they also impinge on classroom behaviour in significant ways. An important and frequent concomitant of the above difficulties is poor educational achievement. The strongest association, however, is with conduct presentations. Thus, Yule & Rutter (1968) report that of 126 eleven-year-old children on the Isle of Wight who showed significant psychiatric disorder, over a third of the conduct disordered children were severely retarded in reading achievement, while reading retardation among neurotic children was little more frequent than in the general population. The questions of whether reading failure leads to conduct problems, or whether conduct problems generate reading difficulties, or whether both are due to some common third set of factors, are still unanswered (Rutter & Yule, 1977).

(iii) Prevalence

With what frequency do the above behaviour problems occur in the child population? Estimates of their prevalence tend to vary. There are a number of reasons for this. First, estimates inevitably reflect the "operational definition" of prevalence that is adopted. Thus:

"An operational definition of a concept...refers to the operations (including any instruments, manipulations, measurements, or recording procedures used in the process of observation) by which the researcher assesses the presence or absence (or magnitude) of the phenomenon denoted by a concept" (Scott & Wertheimer, 1962).

Or, as Wood & Zabel (1978) put it:

"Any incidence estimate is a function of: (a) the person who is estimating, and (b) how he is doing it."

In the absence of agreed definitions of disturbance and universally

accepted measuring instruments, therefore, differences in estimates are inevitable. These may spring, for example, from the adoption of different criteria of severity for inclusion in the category of disturbance, or through varying tolerance levels in the people conducting the ratings - in the classroom context, as Davies (1976) picturesquely puts it: "One man's disruption may approximate another's peak teaching experience."

A second difficulty in estimating prevalence lies in the situation specificity of behaviour. Children may display problem behaviour in one setting such as the home, but not in another, such as the school. Rutter et al., (1970) for example, found that of the 271 children selected.../

selected on the basis of either teacher or parent questionnaires, only 19 were identified on the basis of both. So estimates of prevalence may vary as a function of the number of situations which are investigated. A related aspect of the situation specificity issue is the extent to which disturbance is a function of ecological factors. Thus, surveys conducted in an Inner London Borough and in the semi-rural area of the Isle of Wight have indicated deviance rates on teacher questionnaires and estimates of psychiatric disorder which are roughly twice as high in the city children as in those living in the semi-rural area (Rutter et al., 1970; Rutter et al., 1975). Lavik (1977) presented similar findings in Norway.

A third source of discrepancies is the range of measuring instruments that may be used to identify disturbance and the fact that there may be limited agreement between them. Thus, reliance may be placed on questionnaires or ratings completed by parents or teachers, self-ratings, peer sociometric ratings, behavioural observations, clinical judgment in interviews and so on, either singly or in varying combinations. It is likely that each of these will tap different facets of adjustment and correlations between them may be low (Tindall, 1959; Petrie, 1962). This issue will be considered further in the section on screening procedures below.

It is clear that with so many qualifications attached to their application, prevalence figures must be approached with caution. The most detailed epidemiological studies of behaviour problems in the U.K. have been carried out with 10-year-olds in the Isle of Wight (Rutter et al., 1970) and in an inner London Borough (Rutter et al., 1975).

Children/

Children were identified initially on the basis of teachers' ratings (on the Rutter B Teacher Rating Scale). Subsequently, the children selected on the basis of high questionnaire scores were studied more intensively and compared with a randomly selected control group. Interviews were also conducted with their mothers. On the basis of teacher ratings, 19% and 10% of the children were selected as deviant in the London and Isle of Wight samples respectively. The estimated prevalence rates for psychiatric disorder, based on the more intensive investigations were 25.4% and 6.8% respectively, reflecting a marked regional difference. In a more extensive survey, the National Child Development Study, health, educational and behavioural data have been reported for all children born in one week of March 1958. At the age of seven, each child's teacher completed the Bristol Social Adjustment Guide (Stott, 1963). The results indicated that 14% of the children were rated as maladjusted and 22% as unsettled (Davie, 1972). With older, adolescent children there is perhaps a greater measure of agreement, in this case in estimates of psychiatric disorder. Thus, a follow-up of the Isle of Wight children at age 14 indicated a rate of 20% (Graham & Rutter, 1973), closely matching the 21% figure cited for 13-14 year-olds in an industrial town in the North of England (Leslie, 1974). Similarly, in Norway, Lavik (1977) found marked or severe dysfunction in 19.6% of 15-year-olds in Oslo, compared with 7.9% in a rural area.

In the United States, Glidewell & Swallow (1968) in a comprehensive review of prevalence of emotional disturbance in public schools sponsored by the Joint Commission on the Mental Health of Children in the U.S.A. indicated that approximately 10% of school-age children have emotional problems severe enough to justify special education intervention. Their data additionally suggested subclinical problems of maladjustment in 30%/

30% of elementary school children. State education directors' estimates of prevalence are reported to range from 0.05 to 15% (Schultz et al 1971). More recently, Kelly et al (1977) reported that 2,664 regular class teachers identified 20% of children in kindergarten through grade 12 as having behaviour disorders.

An important feature concealed in overall prevalence figures is that of sex ratio. It appears that, where anxiety-withdrawal problems are concerned, there are roughly equal numbers of boys and girls involved, but more girls develop difficulties in middle to late adolescence. With conduct problems however, there is a clear preponderance of boys, usually in a ratio of about three to one (Graham, 1979). The relative distribution of these two kinds of problems is less clear. In the Isle of Wight study, (Rutter et al., 1970), some 2.5% were classified as having some form of neurotic disturbance, the respective figure for conduct disturbance being 4%.

One thing that is evident from these surveys is the extent to which teachers, who are central figures in many surveys which tap school-based disturbance, identify problems in their charges. Wood and Zubel (1978), in commenting on the American figures, point out that lower estimates of disturbance tend to emanate from school administrators or clinicians who have limited contact with the children. They suggest that the higher rates reported by teachers reflect their needs for assistance in their efforts to cope with stressful situations and promote adquate social development in children. Although the whole spectrum of children's problems might create difficulties for teachers, there is little doubt that the most stressful situations with which they have to/

to deal involve disruptive or disobedient behaviour. There is indeed a longstanding belief (e.g. Uger, 1938) that when identifying problems, teachers tend to focus on children who defy their authority or disrupt the class, and tend to overlook the shy or timid children who do not threaten classroom order. In recent years, there has been increasing concern about rates of disruptive and violent behaviour in schools. As Galloway et al. (1982) point out, it is difficult to establish whether these problems are actually on the increase, since surveys (e.g. Lowenstein, 1975) achieve poor response rates. This, as Galloway et al., suggest, may be because teachers are reluctant to admit that they have disciplinary problems, for fear of looking incompetent in their colleagues' eyes, or because head teachers may be reluctant to admit to outsiders that their school has problems. In a review of these surveys, Laslett (1977), while not playing down the problems of disruptive behaviour, suggested that disruption and violence are not as widespread as the media seem to imply. A similar conclusion was reached by H. M. Inspectors (1979) following a detailed survey of 384 secondary schools, the great majority of which were described as 'orderly communities'. However, indiscipline was a 'considerable problem in 6% of the schools, and a "serious problem" in less than 1%'.

One conclusion that can be drawn from these surveys is that special schools can only cater for a very small minority of those children who would appear to be experiencing problems (Galloway & Goodwin, 1979). Accordingly, as these authors suggest, a massive increase in special educational provision, either within the ordinary school system or outside/

side it, seems to be required. The Warnock Report (D.E.S. 1978) indeed suggested that services should be planned on the assumption that one child in five may require some form of special educational provision in the course of his school career, and one in six at any one time. As outlined in the introductory chapter, the thrust of the present research was very much towards the development of helping programmes within the ordinary school framework, rather than in special settings outside it.

(iv) Persistence of children's problems

Having considered issues in the prevalence of children's problems, it is relevant to ask how long they persist. This is a critical issue for any investigator contemplating treatment of children's problems, given the need to distinguish between treatment effects and the effects of natural developmental progress or improvement. In addition, if such natural or 'unassisted' improvement does occur, is it associated with some types of problems rather than others? The evidence from a number of studies is that many childhood problems are transient and remit without therapeutic help, but that problems of the antisocial/conduct variety are more persistent than their neurotic or anxiety-withdrawal counterparts. Let us examine the evidence for this. In a review of psychotherapy outcome research, Levitt (1957) concluded that the two-thirds improvement rate for treated children was matched in a group of children who received no treatment over a two year period. This untreated group, however, has been criticized as comprising a biased control group (Kolvin et al., 1981; and see Chapter 3). Shepherd et al., (1966) have also presented data pointing to the transience of children's problems. They compared a group of 50 children attending child guidance clinics and a group of children comprising a one in ten sample of children attending ordinary schools in Buckinghamshire, matched

for.../

for age, sex and behaviour. Approximately two-thirds of both groups had improved over a two year period. In a study of 8 - 11 year-olds, Glavin (1967) reported only 30% of a sample originally identified by teachers as still "emotionally disturbed" four years later. An ambitious American longitudinal study in which 6,788 children in second and fourth grades were followed up at two years intervals has been reported by McCaffrey & Cumming (1969). Teachers were asked to identify and describe any children whom they considered to be emotionally disturbed. At the first follow-up in 1963, only 160 children out of 515 identified as disturbed were still rated in this category, and at the second follow-up in 1965, only 126 were again classed as emotionally disturbed. So for the majority of children, their problems were transitory. On a smaller scale, Rubin & Balow (1977) followed a sample of 370 children longitudinally for six years obtaining ratings from their consecutive teachers and an identification as "problem" or "non-problem". Only eleven children (3%) were consistently classified as behaviour problems.

Although many difficulties in childhood may therefore be relatively transient, this is not to say that therapeutic intervention has no part to play. The processes underlying improvement may be accelerated by such help, and the role of support in the alleviation of stress - for the child, his parents or his teacher - cannot be disregarded. A more persuasive argument might be that help should be concentrated on those children whose problems are likely to be persistent, in view of the serious implications for long-term functioning. The evidence of greater persistence for conduct problems than for neurotic problems is fairly consistent. Thus, Rutter et al. (1976) found that three-quarters of children identified/

identified as conduct disturbed at age ten in the Isle of Wight survey, (Rutter et al., 1970) were still handicapped at age fourteen. The comparative figure for emotionally disturbed children was 46%. Similarly, in a five year follow-up of Manhattan children aged five to eighteen, Gersten et al. (1976) found that fighting, conflict with parents and delinquency usually continued or worsened over the period of study. Similar findings are reported in the thirty year follow-up study by Robins (1972) in which children seen in a Children's Guidance Clinic in St. Louis in the 1920's were followed up when they were, on average, 43 years. The findings demonstrated that frequent antisocial behaviour in childhood was a very strong predictor of poor adult outcome and a better predictor than any other factor. The adult outcome for children with emotional disturbance was very much better and few shy, withdrawn or "nervous" children had neurotic disorders in adult life. More recent studies, nevertheless, have suggested that with severe emotional disorders such as persistent school refusal (Tyrer & Tyrer, 1974) or obsessional conditions (Warren, 1965) the prognosis may be poorer. Studies of children who develop delinquent behaviour highlight the onset of serious antisocial behaviour at an early age (Glueck & Glueck, 1964) and point to teachers' and peers' ratings of troublesome behaviour as the best predictors of delinquency (West, 1967; West & Farrington, 1973).

(v) Conclusion

It has been argued in this section that despite the difficulties associated with use of the term 'maladjusted', its retention can be justified, as long as a number of conditions are considered. First, it is suggested that it should be used as an umbrella term requiring closer/

closer and more detailed specification of problem behaviour for individual children. Second, the relational component of 'maladjusted to' compels attention to the context of a child's problems and is a useful corrective to beliefs that maladjustment in some way resides within the child. Third, the assumptions associated with one's use of the term need to be spelled out. In the present context,some of the characteristics of the medical/psychodynamic model which have traditionally been linked with the concept are repudiated in favour of a behavioural framework. It is suggested that usage of the term is not incompatible with such a framework.

In considering the forms that maladjustment might take, the concern of the present study is with those problem behaviours which manifest themselves in the classroom. This allows helping efforts located in that setting to 'reach' their desired target. The incidence of these problems is such that increased or innovative provision is called for. The numbers of children requiring help suggests that the school-based approach of the present study could be one worthwhile option to pursue in the improvement of helping services. Finally, although many children's problems dissipate with little or no help, many of them persist, especially those of the 'acting-out' child. For these problems help may be especially necessary but even with the more transient difficulties, the role of short-term support and help should not be minimized.

(B) The Identification and Assessment of Maladjusted Children

Introduction

Approaches to the identification of maladjusted children have taken a variety of forms. Typical sources of information about children's maladjustment have been ratings by teachers, parents and clinicians, self-ratings by children themselves, sociometric measures, and direct observation of behaviour. These sources may be tapped singly, or in combination. The form taken by identification procedures reflects the use to which information is to be put. Many direct observational procedures, for example, focus closely on specific problem behaviours of interest, such as 'disruptive behaviour' or 'hyperactivity' and the data gathered are often integrated with subsequent treatment interventions. Epidemiological exercises, on the other hand, can be viewed as casting a wider net. These often need to take into account the different settings in which children may show disturbance and the perspectives of different agents or observers, since problems apparent in one setting do not necessarily emerge in others (Rutter et al.,1970; Kolvin et al., 1977).

Although an important part of the present study involved the identification of children with a range of problems, its principal concern was to convey help to such children within the school setting. Such a requirement calls, therefore, for economical and efficient screening procedures which we do not detract from the main intervention exercise. 'Screening' implies brevity in assessment, but in evaluating the complex effects/

effects of treatment, screen measures may not provide adequate information. It was considered necessary in the present study to complement screen data with a number of additional measures - these will be described in Chapter Five along with the screen methods employed. One of the additional forms of assessment employed was direct observation within the classroom. Because of the central position of observational assessment in behaviour modification, the main issues of relevance in this field to the present study will be considered. This will follow an examination of issues in screening which, because of the location of the study, will be concerned with school-based screening. The discussion of screening will include consideration of reliability and validity issues; general concerns in screening within schools; the use of single or multiple measures, and issues in the areas of teacher-, peer-, and child-derived information.

(1) Screening

(i) Issues of reliability and validity.

An important general consideration in screening exercises is that the measures adopted should meet adequate standards of reliability and validity. In other words, where reliability is concerned, a measure should, firstly, produce consistent results when given to the same individual by two or more testers (inter-rater reliability). This might be very relevant, for example, to behavioural ratings on children done by two or more teachers. Although individual teachers' perception might vary considerably, some agreement or convergence would be expected if the ratings are conducted on the basis of extended exposure to chidren in the same setting. A second facet of reliability is test-retest reliability/

reliability or stability. In this case, the same measure given on two occasions should yield similar results. If even short-term stability cannot be assumed, the measure cannot be regarded as useful. With regard to validity, the question here is whether the test actually measures what it is supposed to measure. Validity might be established by showing that the results of the test show some correspondence with results on the same dimension derived from different measurement procedures. Thus, a personality questionnaire tapping social anxiety might be expected to correlate with ratings of behaviour by an observer in an interpersonal situation where anxiety might be elicited. These kinds of agreements can be computed at the same point in time (concurrent validity) or the measure can be correlated with different data gathered at a later date (predictive validity). An example of the latter might be intelligence test scores predicting later exam results.

(ii) Screening within schools

There are a number of important concerns in developing a school-based screen, especially where the screen is not an end in itself but a preliminary to a more elaborate exercise involving treatment interventions within the schools and further follow-up assessments. Disruption of school routine needs to be minimized to avoid interrupting teaching schedules and because of the need to ensure the goodwill and co-operation of school staff following the stage of identification. It is therefore helpful if a screen can be conducted as rapidly and economically as possible, but these considerations should not be permitted to prejudice the screen's efficiency. Teacher ratings, for example, are often vaunted as a rapid and efficient means of screening school-children/

children but the use of this single measure may be, in more than one sense, false economy: first, a punitive rate of mis-identification may result, as discussed below; second, the gathering of teacher ratings may be a more time consuming and tiresome exercise than the collection of self-ratings and sociometric data within classroom groups (Macmillan et al., 1980).

(iii) Single or multiple measures?

School-based screening for maladjustment usually draws on facets of behaviour as seen by the teacher, peers and children themselves. Such a model was advanced over a decade ago by Bower (1969) and it is this model which influenced the design of screening procedures in the present study. There seem to be several advantages in combining these perspectives in a multiple criterion screen method rather than relying on one perspective yielding a single criterion for maladjustment. If 'maladjustment' is taken to connote difficulties in relation to particular demands or expectations as suggested earlier in this chapter, then these have to be taken into account in identification procedures. Thus, within the school framework, for example, children can be regarded as having to meet particular sets of requirements established by the teacher in the sphere of, for example, educational and behavioural functioning, and established by peers in the sphere of social interaction. Problems may be occasioned for some children by the demands of peer interaction but not by those of the teacher. For others, the pattern may be the opposite, pointing to the potential inadequacy of any one measure, whether peer-drived or teacher-derived. Each measure may be seen as reflecting/

reflecting differing vantage points of observers (Strupp & Hadley, 1977) and each one can be viewed as deserving of attention in its own right. To these two vantage points, it is important to add a third - that of the child himself.

When the associations between different measures of adjustment are examined, there is evidence that different domains of functioning and/or different vantage points are being tapped. Tindall (1955), for example, having classified adjustment indices under the five headings of (1) questionnaires and inventories, (2) ratings by adult judges, (3) ratings by peers using sociometric techniques, (4) indices derived from projective techniques and (5) direct observation, found that with a sample of 61 adolescents, the relationships among measures within each of these categories were at a higher level than those between categories. Assessment by one technique appeared to have little value in terms of predicting results from any other technique. Where relationships between teacher, peer and self-ratings were considered, it has been found that correlations are highest between teacher ratings and peer acceptance measures, suggesting a common 'social reaction' (Semler, 1960). The lowest correlations were those between self-report (using the California Test of Personality) and the other two measures. It is important at this stage to examine the contributions and characteristics of these three screen measures of teacher, peer and self ratings, and to offer some justifications for stressing these sources of information in the present study.

(iv) Ratings by teachers.

The use of ratings provided by teachers has had a central role in/

in the identification of maladjustment within the schools (Wood & Zabel, 1978). This is hardly surprising, given the salience of teachers in children's school lives, and much confidence has been placed in teachers' judgments. Thus, Burt (1965) stated that:

"No psychological tests or techniques, no interview by a psychologist or a clinical psychologist, can yield such trustworthy assessments or such sound predictions as those of an experienced teacher who has watched and studied his pupils' behaviour and development, year after year.' (p.375)

This is indeed a strong expression of faith, and underlines the important contribution that teachers can make in screening exercises. However, the value of this contribution can be overstated. It is instructive to recall Guildford et al.'s (1962) suggestion that analysis of ratings is useful primarily to discover what is going on in the minds of the raters and should not be assumed to be valid in the sense that they provide generalised descriptions of personality traits. The ratings provided by teachers will inevitably reflect the manner in which they structure their perceptions of the classroom and children's behaviour therein. For example, one might expect teachers to emphasise those aspects of behaviour which have direct implications for control and management of the class. Such is the inference to be drawn from Wickman's (1928) early study in which teachers, by comparison with 'mental hygienists' (psychiatrists, psychologists, social workers), were found to stress the importance of misbehaviour in the classroom, with problems indicative of withdrawing, recessive characteristics being seen as of comparatively little significance. Despite the dubious methodology of this study (Watson, 1933), Eysenck & Cookson (1969) suggest that this bias has been..../

been confirmed in many subsequent investigations. Concluding a review of the field, Hollins (1955) suggests that:

'A change of approach may make a difference as to whether teachers regard violations of morality or school offences as the most serious forms of problem behaviour, but there is overwhelming evidence that they do not regard the introverted, unsocial child as a problem.' (p.22)

Garner & Bing (1973) have also suggested that teachers may overlook quiet, passive, but potentially disturbed children.

Another salient dimension in teachers' ratings has been highlighted by Hallworth (1966) and Cookson (1973) who demonstrated that teachers' notions of stability in children are heavily linked up with their general approval or disapproval. In evaluating emotional stability, therefore, teachers may, in effect, be asking themselves, "How much do I approve of him?" or "Do I consider him a good pupil?" Tsoi and Nicholson (1982) report that teachers' ratings tended to be factorially inter-related, suggesting that rather than rating along discrete dimensions, they may rate according to a global continuum, like good versus bad. Teachers also seemed to be more consistent when rating items that were either of a disciplinary nature or else had to do with a child's level of activity. It has been suggested by Herbert (1974), however, that this overriding 'approval' constraint may tend to operate most when teachers are asked to rate 'stability' in a global fashion. His own work suggests that if a detailed set of behaviours is provided, teachers can apply a much more sensitive and differentiated conceptual scheme. In general, however, a more comprehensive view of adjustment may be obtained by supplementing teachers' ratings by information provided by the children themselves.

There are additional reasons for.../

for not relying on teachers' ratings alone. First, teachers may not be sufficiently sensitised to the interpersonal difficulties experienced by some children where these do not have 'visibility' in the teaching contact. Sociometric information can be a useful addition here. Second, as the sole criterion of maladjustment, they may be inefficient, in terms of misclassifying children as false positives ('normal' children classified as 'maladjusted') or false negatives ('maladjusted' children classified as 'normal'). Mulligan & Douglas (1963) had teachers complete questionnaires in respect of 13-year-old children attending child guidance clinics and who had been diagnosed by psychiatrists as suffering from aggressive or neurotic disorder. They compared ratings on these children with those provided for 'normal' matched controls. This produced a false negative rate of 24% and a false positive rate of 29%. In Rutter et al.'s (1970) Isle of Wight study, 157 children were selected by teacher questionnaire as 'maladjusted'. In this study, children were additionally selected by a parent questionnaire, and a final diagnosis was arrived at in respect of each child on the basis of interview data gathered by a child psychiatrist and much additional information obtained from the parents. Of the 157 children selected on the teacher questionnaire, 64 were finally diagnosed as maladjusted - a false positive rate of 59%. A similar screen by Rutter et al. (1975) in an Inner London Borough, this time employing the single criterion of teacher rating, yielded a 57% false positive rate. Macmillan et al. (1980) compared the efficiency of this single screen exercise with that of a multiple-criterion screen, involving teacher, peer and self ratings, with a comparable population. They estimated the rate of disturbance by/

by the validating method of rating by a psychiatrist as to the presence of disorder. The multiple criterion screen selected a smaller number of cases of whom 78% proved to be disturbed on clinical rating and the single criterion screen selected a larger number of cases of whom 43% proved to be disturbed clinically.

These criticisms of teacher ratings as the only indices of maladjustment are, of course, predicated on the acceptance of psychiatric judgments as valid criteria for ascertaining 'maladjustment' or disturbance. The criterion for 'efficiency' here may be a harsh one, in that clinical ratings of adjustment may have a generality of reference which teacher ratings clearly do not have. Indeed it may be argued that if one is concerned with problem 'behaviour' rather than with 'adjustment', little further validation is required for a teacher's ratings. If they indicate that child X presents difficult behaviour for teacher Y, then, for this teacher, that is the reality which has to be confronted. If one is concerned only with this limited situation then a divergence of perceptions between teacher and clinician is not of fundamental importance. If a more general statement concerning children's adjustment is sought, teachers' ratings cannot be expected to provide all the necessary data, and additional sources must be tapped. The treatment interventions in the present study were <u>not</u> guided solely by teachers' perceptions, so supplementary sources were pursued.

<u>(v) Ratings by peers - sociometry</u>.

Sociometric measures are the most commonly used methods of identifying children with difficulties in social or interpersonal relationships, and/

and of establishing a child's social status relative to other children in a group. The companionship choice technique developed by Moreno is perhaps the most common of these. Typically children are asked to name their best friend(s) and/or the peer with whom they would most like to play, work or sit beside. The child's sociometric score is then computed based on the number of nominations received. Such an approach allows identification of popular children, or sociometric 'stars' - those receiving most choices - and of 'isolates' - those receiving least.

A number of issues require consideration in designing a sociometric measure. First, researchers differ as to preference for employing specific criteria such as the above, which refer to concrete situations, or more general criteria, e.g. asking the question, "Who is your best friend?" Moreno (1953) asserts that specific criteria are likely to yield the more reliable and valid data. Others suggest that scores based on a more general criterion are likely to be much more stable, because the variability associated with specific situational factors is eliminated (Gronlund, 1955; Harper, 1968). Since there is some evidence of a general factor of acceptability underlying sociometric choices (Frankel, 1946; Bjerstedt, 1956) it may make little difference which method is used.

A second concern in the application of sociometry is the question of how to deal with the fact that the number of nominations a child receives depends on the number of raters making an assessment. The smaller a class group, for example, the fewer choices there are to be spread around. Another potential constraint stems from the sex distribution within the group, where cross-sex choices tend to be avoided. Bronfenbrenner/

Bronfenbrenner (1943) described a form of ready-reckoner to make the necessary corrections. This employs a chance model, identifying persons who receive significantly more or fewer choices than would be expected by chance. Adjustments can be made for variations in class size, so that comparability can be achieved in terms of number of possible choices that can be received. Macmillan et al. (1978) found that with class size ranges of 24-34 children, the correction made very little difference to raw scores, both where correction was made for class size and for the combination or separation of boys and girls. Raw scores could therefore be used with groups in this range.

A third issue that requires consideration is whether negative nominations should be employed, in addition to the customary approaches to identifying popular and unpopular children. The use of negative nominations is intended to differentiate isolates, who may merely be ignored by their peers, from those who are actively disliked or rejected. While some unfortunate children are both isolated and rejected, these two sociometric characteristics do not necessarily go together. They are not merely different labels for the extreme end of the one dimension (Sells & Roff, 1967). Only moderate negative correlations are typically obtained between positive and negative choices (Hartup, 1970; Macmillan et al., 1978). With a negative nomination format, children are usually asked who they would <u>not</u> like to play with/work with, etc. There has been a tendency in sociometric studies to emphasise positive rather than negative choices, for a number of reasons. Possibly the major reason is the objection that the use of rejection choices will draw undesirable attention to those children who are negatively evaluated and encourage the crystallisation of negative feelings towards them.

It/

It is also suggested that rejection measures may evoke resistance and resentment on the part of either children or teachers since our culture does not favour admission of negative feelings (Harper, 1968). A further objection is that rejection is an artificial measure since people are not interested in those with whom they do not associate (Northway, 1952). However, schools do impose a framework on children's interaction and they are frequently obliged to associated with peers they might prefer to avoid.

Despite these potential difficulties, there is little in the sociometric literature to point to negative effects of using rejection measures (Foster & Ritchey, 1979). If one is concerned with the identification of children who are having problems of social adjustment - as in the present study - then the discovery of rejected children would seem very pertinent. They are perhaps more likely than their isolated counterparts to be miserable and unhappy, though this should not be assumed automatically from indices of sociometric status. Inferences such as these need to be made on the basis of more comprehensive information than that supplied by sociometry alone. Certain isolated children, for example, are not adversely affected by being relatively friendless: Gronlund (1970) described 'self-sufficient' children who deliberately withdrew from interaction and did not seek their peers' attention.

These, then, are some of the problems to be dealt with in designing sociometric measures. What is the validity of the information gathered by such measures? Despite caveats such as Gronlund's, there is a substantial body of evidence showing that unfavourable sociometric status tends to correlate with indices of maladjustment. A number of studies employing/

employing scores on the California Test of Personality as adjustment criteria have shown a close relationship between lack of popularity and poor psychological adjustment: this was true of elementary (Bonney, 1944) and secondary school age children (Bauer, 1971; Grossman & Wrighter 1948; Semler, 1960). Self-reported anxiety - for example as measured by the Children's Manifest Anxiety Scale - similarly appears to be associated with poor status amongst peers (Horowitz, 1962; McCandless et al., 1956; Trent, 1957). The same pattern of relationships is evident when adjustment is judged by teacher ratings (Glidewell & Swallow, 1968; Yellot et al., 1969; Chazan, 1963; Ullman, 1952).

Numerous other characteristics indicative of psychological difficulties have been associated with poor sociometric status, such as, for example, the presence of psychosomatic ailments (Izard, 1959), social immaturity (Shaw, 1952), aggressive behaviour (Hartup et al., 1967) and delinquency (Croft and Grygier, 1956).

In addition to these cross-sectional relationships, sociometric indices also have clear predictive utility. Thus, unpopular children are more likely to be disproportionately represented later in life in a community-wide psychiatric register (Cowen et al., 1973). Amongst the indicators identified in the Austin Longitudinal Research Project (Currie et al., 1974) as precursors of inadequate personal adjustment some nine years later, rejection by peers in the first grade proved to be one of the most useful. In a four year longitudinal study of 4,000 children, Roff, Sells and Golden (1972) found that except for the lowest socio-economic class, the relationship was highly positive between delinquency and low peer acceptance scores taken four years earlier/

earlier. Delinquent behaviour has been similarly predicted by Gibson & Hansen (1969), Harper (1965) and Skaberne et al. (1965). Such relationships are of course correlational, and should not be taken to mean that poor peer relations cause later problems.

From this brief review, it is apparent that sociometry is capable of yielding information that is rich in implication for children's current and future adjustment. What sociometric assessment does not do is establish the nature of a child's interpersonal problems. It lends itself to classification rather than specifying deficits and assets around which an intervention programme can be structured. Nevertheless, where intervention is intended, sociometry provides an efficient means of identifying children requiring help, which can then be followed by a more detailed and focused analysis.

(vi) Ratings by children.

An important consideration in the choice of screening procedures in the present study was the provision of an opportunity for children to make self-statements concerning aspects of their adjustment. This was on the assumption that self-ratings by the child serve to complement teacher- and peer-derived information and may reveal the personal unhappiness and unease which these other sources may bypass. Ratings by teachers and peers perhaps reflect the more 'visible' aspects of a child's problems, rather than hidden worries, anxieties or feeling-states.

There is some evidence for these assumptions. The possible lack of overlap that may be found between self-ratings and information from other sources (e.g. Tindall, 1955; Semler, 1960) has already been considered. It is apparent that, in addition to other possible limitations/

limitations discussed above, teachers' ratings may not show a close relationship with descriptions the children may provide of themselves in the sphere of adjustment or emotional stability. Nicholson & Gray (1972), for example, compared ratings by teachers on a scale containing adjectives that described the dimensions on the Junior Eysenck Personality Inventory (JEPI) (Eysenck, 1965). Adjectives in the teacher scale for the neuroticism dimension, which reflects emotional stability, included: moody, anxious, rigid, touchy, restless calm, reliable, carefree and stable. Strong negative correlations were found between JEPI neuroticism and teacher-rated measures on this dimension. Similar discrepant results have been noted in other investigations by Eysenck & Pickup (1968), Eysenck & Cookson (1969) and Cookson (1973). More recently, in an investigation of teacher ratings, Tsoi & Nicholson (1982) suggested that neuroticism is the construct teachers find most difficult to rate. In a factor analysis of teacher ratings of child behaviour, this factor took up only 2.5% of the variance. In addition, the ratings on this dimension did not correspond to the ratings of neuroticism obtained on the JEPI.

There are a number of concerns associated with the use of self-ratings of questionnaire response. First, the child may want to present himself in the best light and may wish to give socially desirable responses (Edwards, 1957). Second, clear-cut responses may be difficult to give, and some children may be unhappy about the forced choices which some questionnaires require. Third, the fluidity of meaning in many words or questions make it difficult to be sure that they/

they mean the same thing to any two respondents. Fourth, it is important to establish that children can read and write adequately if they are to complete questionnaires. Where the identification of maladjusted children is concerned, this is an important point, since there is a far greater incidence of maladjustment in the backward classes of a school than in the other classes (Chazan, 1963).

A final concern is where these questionnaire responses are cast in the form of personality assessment. Traditionally, such assessments have entailed assumptions in which personality has been viewed as a reflection of enduring states or traits which underlie behaviour and which are expected to be consistent across settings (Hartmann et al., 1977). This perspective has come under attack in recent years, as evidence has grown concerning the situation-specificity of behaviour, and the extent to which an explanation of behaviour needs to be approached through consideration of person-environment interaction (Mischel, 1973; Bowers, 1973). The contribution of personality tests may therefore need to be viewed as more circumscribed than before. As Moore (1974) comments: "It is not surprising that no test ... has been found to have very high validity for the prediction of behaviour tendencies which cannot themselves be shown to have permanent existence; or for assigning subjects unequivocally to category or personality types too unsubtle to encompass the variability of an individual" (p.129). But such measures are not therefore to be considered valueless. For all their possible shortcomings, they still constitute an important set/

set of self-statements in their own right; they can be viewed as samples of behaviour, within a framework which stresses overt behaviour; and amongst the array of possible measurement approaches, they are not exclusively prone to problems of reliability and validity. Finally, the question of their value is largely an empirical one: the inclusion of such measures in a preliminary screen battery is justified where their contribution is weighed and assessed against a background of other data sources, derived both from the screen itself and from more intensive subsequent assessments.

(vii) Conclusion.

What implications may be drawn from the foregoing discussion of screening procedures for the present study? First, a school-based screen should be as economical as possible, so as to create the minimum disruption to the regular timetable and so as not to prejudice the co-operation of staff, both in the initial screening exercise and in any subsequent assessments that may be required. This should not, however, place significant limits on the screen's efficiency. Second, screen measures should be sought which can be applied with acceptable standards of reliability and validity. Third, multiple screen procedures would seem preferable to reliance on any single measure. Within the school context, teachers, peers and the child himself represent three important vantage points from which to view a child's adjustment. Information should be sought from these three sources, with the sources viewed as complementing one another. This broad perspective on problems of adjustment is an important consideration for the present study, since the intended treatments were geared to intervene with a wide spectrum of problems within the schools, rather than focusing on one or two limited aspects of maladjustment.

(2) Observational Assessment

(i) Introduction.

Observational procedures were employed in the present study as a means of focusing directly on aspects of children's and teachers' behaviour in the classroom. These were aspects which were of central concern in the treatment intervention and the provision of observational data was an attempt to monitor the extent to which changes occurred. Reliance on observational assessment has been a principal feature of behavioural approaches and the development of behaviour modification is marked by a rejection of traditional forms of indirect assessment in favour of measurement of actual behaviour. However, it has become evident in recent years that there is no simple purity in observing and recording behaviour which ensures intrinsic validity, that many of the problems associated with traditional forms of assessment are not avoided, and that the whole enterprise is exceedingly complicated (O'Leary & Johnston, 1979). This section will briefly highlight the main issues and problems of concern to the present study. These will include: the methods of assessment; the units of observation and their definition; the need for training of observers; issues related to inter-observer agreement and to validity, and a number of sources of error and bias. Extensive coverage of the area is available in Johnson & Bolstad (1973), Kent & Foster (1977), and Kazdin (1978).

(ii) Methods of assessment.

The methods of assessment refer to how observed behaviour is scored and converted into data. The most commonly used methods in classroom behaviour/

behaviour modification have been frequency and interval recording. Other approaches involve duration measures - recording how long a response lasts - and counting the number of individuals who perform a response.

With frequency measures, the observer simply tallies the occurences of particular responses. It is particularly useful with discrete responses with a clearly definable beginning and end. Ongoing behaviour, however, such as smiling, talking and studying, are not easily handled by frequency measures because they are not sufficiently discrete and their duration varies so widely. Despite the simplicity of frequency measures, the difficulty presented by continuous responses is an important consideration for many classroom studies, where measurement of attention to classroom tasks - amongst other behaviours - is often a particular concern. An interval recording format is appropriate here. In this method, behaviour is recorded during a short block of time, such as thirty minutes or an hour, once a day. This block of time is divided into shorter intervals (e.g. ten second intervals). During each of these intervals, the target response is recorded as having occurred or not occurred. Several responses of the same type within the one interval are not counted separately. If the behaviour occurs continuously across several intervals, it is recorded in each one in which it occurs. In a variation of interval recording, referred to as time-sampling, behaviour is observed at several separate moments in time (e.g. every fifteen minutes, or once an hour) rather than in periods in immediate succession.

Kazdin.../

Kazdin (1981c) suggests that interval recording is probably the most frequently employed strategy for three reasons. First, because it is flexible, any observable response can be recorded with it. Second, several different responses can be scored simultaneously, for a given interval. Third, the results of interval scoring are easily communicated to others. A percentage measure of the 'amount' of the response can be computed, based upon the proportion of intervals in which it occurs. Interval recording is consistent with the goals of many programmes that attempt to increase the time over which responses such as studying, talking or socialising occur. Percentage measures reflect this goal directly.

(iii) Units of observation and their definition

The question of defining the units in which behaviour is to be recorded is an important one. There are countless ways in which the 'stream' of behaviour can be broken up and analysed. What constitutes a meaningful unit or 'chunk' for one investigator may not be seen in the same light by another. Connolly (1975) points out that 'observing behaviour...raises many problems, the most obvious of which is not knowing initially at least, what to measure. No observer is totally free of bias and inevitably certain features of the total situation will be selected while others are largely ignored. Also there are limits on the amount of information the observer can process and record per unit of time.' (p.75) Some researchers try to deal with units of behaviour which go together in natural blocks, as in interaction sequences, but as Brinich (1981) points out, options vary as to what these 'natural' units might be.

Becker.../

Becker et al. (1967) were amongst the first to develop discrete, definitive categories of child and teacher behaviour and their categories have been used by many other researchers with minor variations and additions. They suggested the following rules for establishing categories for children's behaviour: (1) They should reflect behaviours which interfere with classroom learning and/or (2) they should involve behaviours which violated classroom rules and/or (3) they should reflect particular behaviours a teacher wants to change, (4) the classes should be constituted by behaviours which are topographically similar in some important way, (5) the classes should be mutually exclusive, (6) the definitions should refer to observables and not involve inferences and (7) the number of classes should not exceed ten. These authors also employed a coding system in which child and teacher behaviours were recorded jointly, so that possible functional relationships between the two could be explored. Such data could usefully inform treatment interventions.

Considerable care is necessary in the preparation of adequate definitions for the behavioural categories to be addressed. A recent example of the fine distinctions that might be necessary for some investigative purposes is the work of Arlin & Roth (1978). They were interested in the relationship between time-on-task and reading gains and hypothesized that time-on-task would be important as a determinant of reading growth only to the extent that a pupil is <u>actually reading</u> during.../

during the time-on-task. A distinction was therefore necessary between time-on-task and time-on-reading. They write:

'Although the task at first seemed quite formidable, with practice certain signs became quite clear indicators of reading or not reading. Many of the pupils at this age still moved their lips while reading or placed their fingers under the words. We took this as a sign of on-reading. We took as off-reading signs such as flipping through the pages too quickly (we assumed no speed reading capabilities)... Eye movements and head movements would indicate that the pupil was glancing at random over the page rather than reading systematically from upper left to lower right.' (p.207)

This illustrates the very demanding observational requirements when a fine-grained analysis of behaviour is sought.

An issue that has been addressed little in the literature on behavioural observation is the demands that are placed on the observer operating a multi-category system over a prolonged period of time. Strict and sustained vigilance is called for in these situations and effects of fatigue probably contribute to error rates in no small measure.

(iv) Observer training

It is evident that, given such demands on the observer, in the sense of vigilance on the one hand, and on the other the need to pay close attention to definitional requirements, training in the use of the observational system is a prerequisite for any research exercise. This training should include: learning the operational definitions and scoring procedures of the observational system; discussion of procedural problems; in situ practice in settings similar to the research settings; and 'recalibration' sessions where definitions and their appliction are re-checked (Hartmann & Wood, 1982).

(v) Computing inter-observer agreement.../

(v) Computing inter-observer agreement

In typical behavioural observations, an observer scores an individual's behaviour according to a pre-specified code. Scoring behaviour requires that an observer make judgments about the occurrence of the response. In the fast flow of interactions in a social setting this is often no easy matter, and thus the extent to which behaviour is observed and scored consistently and reliably is an important issue. Good inter-observer agreement, or reliability, is central in assessing and evaluating behaviour changes. Agreement can be seen as reflecting the clarity of response definitions as well as accuracy of observation.

There has been considerable discussion in the literature concerning issues in the computing of observer agreement, and of the different methods to be used (e.g. Hartmann, 1977; Frick & Semmel, 1978; Kazdin, 1979). We are concerned here with methods for computing agreement for interval recording, and these typically take the form of percentage agreement or product-moment correlation. Percentage agreement is computed as the number of intervals in which agreement occurs between two observers divided by the total number of agreements plus disagreements, multiplied by 100. As noted by Hartmann (1977) this method has its appeal primarily in its computational and interpretative simplicity but it unfortunately has limitations. An important concern is that percentage agreement is affected by the frequency of occurrence of the behaviour observed. A positive relationship has been observed between frequency of behaviour and per cent agreement such that low and high frequencies are associated with lower and higher observer agreements.../

agreements respectively (Johnson & Bolstad, 1973). The problem here is that, as the frequency of a behaviour recorded for a session increases, so does the number of agreements expected by chance, so a correction for chance agreements may be necessary.

Where correlation coefficients are computed as a reliability index, the observers' totals for intervals of agreement are correlated. The reliance on totals, however, bears with it the problem that two observers could correlate very highly even though they did not agree on a given occurrence of the response. As Kazdin (1978) points out, correlation hides the moment-to-moment level of agreement and because this method allows an extremely large discrepancy between observers, it is used infrequently in calculating reliability. Most writers therefore recommend the percentage agreement formula, with attention being paid to the need for chance corrections.

(vi) Establishing the validity of behavioural observation

As pointed out by O'Leary & Johnson (1979), it must be demonstrated that observational data are indeed valid and not simply assumed to be so. Is there evidence that behavioural observations correlate with other dissimilar methods of measuring the same variable? This, of course, is a question for specific observational systems and what evidence is available is not to be regarded as generally applicable. Some examples are studies by Forness & Esveldt (1975) and Forness et al. (1975) which show that teachers' ratings of children's classroom performance correlated with observational measures, with the most agreement occurring for extreme groups. Predictive validity was demonstrated by/

by Cobb (1972), with behavioural observations taken during arithmetic lessons correlating from 0.63 to 0.69 with standardised mathematics achievement test scores. Similarly, McKinney et al. (1975) found correlations of 0.60 between behaviour ratings in the autumn term and achievement in the following spring.

(vii) Some sources of error in direct observation procedures

A number of possible sources of error in conducting observational exercises have been identified which may affect the nature and quality of the data. Two major difficulties are observer bias and observer reactivity. In the first, the observer's biases or expectations about the behaviour being observed may influence the recording of observations. In the second the presence of an observer may change or affect the actual behaviour being observed. This latter effect may operate not only at the level of, for example, observer and child, or observer and teacher, but also at a different level, where observer reliability exercises are conducted when the observer himself is observed. These biases will be examined in turn.

(viii) Observer bias and expectancies.

The issue of experimenter bias is not a new one in psychology and was perhaps made most salient by Rosenthal in his controversial report which claimed to show that the beliefs of teachers affected the intelligence test performance of their pupils (Rosenthal & Jacobsen, 1966). In the context of direct observation, distortions can be produced by the expectations held by the observer. However, these distortions tend to operate at the level of global evaluations or subjective/

subjective impressions given by observers, rather than at the level of behaviour ratings per se (Kent et al., 1974; Shuller & McNamara, 1976). In the Kent et al. study, some observers were told that children's disruptive behaviour (on a video tape) would decrease, others that it would not change from baseline. The data on video tape in fact showed no changes across phases. Over all, expectancies did not influence observer recordings, but when observers were asked to characterise the effect of the programmes on a questionnaire, their evaluation reflected the expectancy they had been given. Siegel et al. (1976) found that neither behavioural observations nor rating scale responses were affected by differential expectations, and they attribute this to the precision of definition of behavioural items in both forms of recording. Clearly, the less room for inference, the less opportunity there is for interpretations to stray in the direction of expectancies.

In an interesting study with important implications for research interventions, O'Leary et al., (1975) informed observers of a video tape that disruptive behaviour would alter from baseline levels in response to a token programme. Actually, tapes for the baseline and treatment phase were matched for disruptive behaviour. Positive comments were provided by the experimenter if a reduction in disruptive behaviour was recorded by the observers, and negative comments (disappointment) if no change or an increase was scored. This combination of an expectancy to see change, and 'shaping' of recordings by the experimenter, was found to produce consistent bias in behavioural frequencies.

(ix)/

(ix) Reactivity of observational procedures

This source of error refers to the possible effects on those observed of having a person present who is monitoring their behaviour. If simply observing behaviour serves to change that behaviour one cannot safely generalise from the observational records to situations in which observers are not present. The nature of the effects on observed subjects, however, is not clear. Surratt et al., (1969) suggested that the observer's presence led to more 'on-task' behaviour in the classroom, but Dubey et al., (1977) found no systematic effects of observer presence or absence on the classroom behaviour of children or their teacher. Mercatoris & Craighead (1974), with a small class group of six retarded children found that observer presence did increase the frequency of teacher-pupil interactions, but not the appropriateness of pupil behaviour. Samph (1976) found that teachers' verbal behaviour was more like their perceived ideals when there was an observer present.

The evidence on this important issue is unfortunately not extensive and firm conclusions cannot be reached as to the likely effects on observed subjects. An obviously important question is whether any reactivity - if it exists - actually persists for a long period of time, but it is a question which has not been directly addressed in classroom settings. One would hope that any reactive effects are short-lived as children and teachers habituate to the presence of an observer. Time may need to be devoted to such an habituation period at the outset of observation, but as Samph (1976) points/

points out, there is little guidance as to how long this should be.

(x) Biases in the assessment of inter-observer agreement.

The above influences which flow from observer expectancies or from observer presence also play a part when checks are made on the agreement between observers. This might involve having two observers recording the same behaviour, or having an experimenter present in a 'supervisory' capacity to monitor the observer's recording. In whichever format is used, the observer is now himself being observed. A number of studies have indicated that when individuals are aware that agreement is being assessed, inter-observer agreement is higher than when they are unaware (Reid, 1970; Romanczyk et al., 1973; Kent et al., 1977). The implications of such a finding for the quality of data are important, particularly when one considers Romanczyk et al.'s finding that when unaware of reliability checking, observers recorded 20% less disruptive student behaviour.

The consistent pairing of specific observers to conduct reliability checks may raise further problems in interpreting their agreement. Evidence suggests that observers who consistently work with each other and receive feedback on each other's performance develop idiosyncratic variations of the original response definition, so that behavioural codes are altered over time (O'Leary & Kent, 1973). This phenomenon is referred to as observer 'drift', or 'instrument decay', and it can obviously occur in routine observation as well as in the context of agreement checking. As well as referring to gradual change in definitions of behaviour, instrument decay also covers/

covers deterioration in measurement occasioned by, for example, missing episodes of behaviour, adopting short cuts by collecting only parts of the data and so on. Many of these problems can be countered by arranging for random or covert checks (Taplin & Reid, 1973). The need for clear and precise behavioural definitions has already been mentioned as a means of reducing expectancy biases: it is also an important requirement in facilitating observer agreement.

(xi) Concluding note.

A number of recommendations for the present study can be drawn from this brief review of research on observational methods, although it should be pointed out that most of the work on expectancy and reactivity effects has appeared since the study was conducted. First, the general complexity of the task of observation, with its demands for strict and sustained vigilance, suggests that a good deal of training is required for observers. It is surprising that little attention has been paid to the effects of fatigue in such exercises, given its likely impact on error rates. Second, there is a clear need for precise and unambiguous behavioural definitions, which may have important implications both for the reliability of the system and for the reduction of expectancy effects. Third, the extent of inter-observer agreement requires to be checked, ideally under covert and unobtrusive conditions. Fourth, given that knowledge of treatment objectives may influence recordings, it is important, from the 'experimenter's' point of view, that discussions with the observer do not involve comments concerning the desirability or otherwise of recorded changes. Finally, even although the evidence concerning the effects of observer presence on children's behaviour is unclear, it may be wise to allow time for a class to habituate to the observer's presence.

Having examined some of the conceptual issues in the concept of 'maladjustment', and issues in the identification and assessment of maladjusted children, the next chapter will address the theoretical underpinnings of behaviour modification as one approach to helping such children.

CHAPTER 2

THE THEORETICAL BASIS FOR BEHAVIOURAL TREATMENT PROCEDURES

THE THEORETICAL BASIS FOR BEHAVIOURAL TREATMENT PROCEDURES

Introduction

This chapter will provide a description of the theoretical bases for the treatment procedures in the present study. These procedures were grounded, for the most part, in principles of operant behaviour, developed and elaborated in tightly controlled laboratory settings. An understanding of these underlying principles is important for two main reasons. First, behavioural interventions with children are not based solely on positive reinforcement (or reward) but often on diverse procedures, sometimes used in combination. Paying attention to the variety of principles involved is therefore necessary. Second, the effectiveness of the procedures employed is crucially dependent on the manner in which they are applied, and as indicated in this chapter, there are a number of important dimensions of application which must be attended to. The translation of these laboratory-derived procedures into the less well-controlled conditions that commonly prevail in 'treatment' or 'applied' settings often creates some difficulties and some of the issues here will be addressed. Before describing these procedures, however, and some of the possible problems of transposition outside the laboratory, it is appropriate to describe the philosophical position in which they are embedded. This will help to clarify what the position asserts in its approach to understanding behaviour, and what its proponents view as its strengths. It will also provide a context for understanding what some of its critics view as its weaknesses - these criticisms will be considered later in the Discussion Chapter.

(A) The Philosophical Basis of the Operant Approach

Blackman (1981) identifies the use of operant conditioning techniques as one of three logically separate aspects of an approach within psychology known as the experimental analysis of behaviour. The other two threads are (i) the search for empirical statements about behaviour which are valid and reliable but which relate to individuals rather than to differences between groups and (ii) the broad philosophical position of radical behaviourism. (Blackman acknowledges that despite the exclusiveness implied by the title of 'the experimental analysis of behaviour', it is merely one approach to investigating behaviour experimentally.) Both these related threads have important implications for the present study and each will be examined in turn.

(i) The emphasis on the study of individuals

The first of the above themes is of particular importance since the present investigation, while employing procedures developed in the context of single-subject research, applied them within the framework of a group design. Some clarification of the kinds of statements which may flow from these two kinds of design is offered by Bakan (1967). He distinguishes between aggregate propositions and general propositions. The former assert something which is presumed to be true of samples of subjects considered as aggregates, while the latter assert something which is presumed to be true of each and every member of the sample. Group approaches often lead to statements about differences between samples on the basis of differences in average scores. These statements may be of an aggregate/

aggregate rather than a general form, i.e. the differences may not necessarily relate to all the subjects. Bakan notes that psychologists often, on the basis of studies of groups, shift to unwarranted statements about individuals on the basis of aggregate and not general statements. (Group comparisons may also, of course, permit general statements.)

Workers within the tradition of the experimental analysis of behaviour aim explicitly for general propositions which are true of all members of a designable class. They approach such propositions through studies of a limited number of individual subjects. A number of experimental designs have been generated for the analysis of within-subject effects in such studies (Gelfand & Hartmann, 1975). The most commonly used designs are reversal or ABAB designs, and multiple-baseline procedures. The basic logic of these designs is to determine operations that relate functionally to the performance of behaviour. In a reversal design, the effect of a variable (e.g. food as a reinforcement) is demonstrated by following a base-line period of observation with consecutive phases of presentation, withdrawal and re-presentation of the variable. Control over the behaviour is demonstrated if it is altered predictably by the experimental operations. The multiple-baseline procedure involves establishing baselines for several behaviours that occur concurrently in the experimental situation. Reinforcement is applied to occurrences of only one of these behaviours. A change in this behaviour and no change in the concurrent behaviours is accepted as a demonstration of the specific effect of the contingency. The reliability of the effect is further established by subsequently reinforcing other behaviours for which baselines have been established and noting changes in those behaviours.

Despite/

Despite the emphasis within this tradition on individually oriented strategies of analysis and experimental control, there is no logical reason why operant conditioning procedures cannot be employed within group designs which may lead to aggregate propositions via statistical comparisons (Blackman, 1981). There are indeed a number of situations in which single-subject approaches are inadequate and groups designs are relevant. Thus, where the relative effectiveness of different treatment approaches needs to be explored in applied settings, group designs are necessary (Kazdin, 1973). Similarly, where questions relating to the long-term effectiveness of a treatment need to be addressed, group designs, with appropriate controls, are recommended (O'Leary & Kent, 1973). With a view as to how evaluative procedures can make maximum impact on figures such as politicians, educators and administrators, O'Leary & Kent (1973) suggest that a preoccupation with single-subject designs to the exclusion of other designs may constitute a weakness rather than a strength. They suggest that those who do not share the conceptual framework of the behavioural approach may view the ready reversibility of effects as evidence of the transience or triviality of the behaviours being dealt with. A more damaging criticism they offer is that concern with reversibility forces investigators to focus on limited behaviours, which may be readily influenced but of no great practical significance. However, it needs to be pointed out that adoption of group designs and evaluation by Fisherian statistics does not solve the problem of communication with policy makers (Kiernan, 1974), nor does statistically significant effects with large numbers of subjects guarantee treatment effects of genuine value./

value. Clearly, different research questions require different methodologies and designs, and single-subject and group designs should be seen as complementing one another. Operant and non-operant designs each have important contributions and in the words of Hogg (1976):

'however convincing the internal adequacy of a behaviour modification programme's success may seem to its proponents, it is questionable whether such isolationism can survive indefinitely.' (p. 175)

The above drawbacks of individual subject designs were important considerations in determining the organisation of the present study in the form of a group design. Indeed, the treatment comparison and long-term follow-up requirements effectively precluded an intra-subject design. Other important considerations were that reversal procedures may be viewed as quite inappropriate in an applied setting and may indeed not be attainable. Thus, if control has been established over undesired behaviour (e.g. disruption in a classroom) a move to, in effect, reinstate it for purposes of demonstrating experimental control may not be well received by people who are exposed to it. In some situations it may be difficult to effect a reversal and return to exactly the same baseline conditions after an experimental programme has been applied. Classroom rules, for example, may be difficult to reverse in a credible fashion, or a teacher may not find it easy to modify her own experimentally manipulated attention in line with design requirements.

The intensive analysis of individual cases, with painstaking regard for experimental control is clearly a major strength of the operant/

operant approach and the philosophical influences which inform it. There is accordingly a risk that in applying its procedures within a framework (such as group designs) which would be considered alien, these strengths may be diluted. These risks may have implications for both quality of design and quality of treatment intervention. Thus the information that may be sought in a clinical context for individuals may not be readily available from a study of groups of subjects. The more subjects we employ, the more remote from individuals we may become. From the treatment perspective, there is the danger that a simplistic technique-oriented approach is adopted without the necessary regard for environmental and functional analysis. However, as Franks & Wilson (1979) have observed, it is often necessary to trade off the advantages of sophisticated methodology against the exigencies of the applied setting. 'To preserve the integrity of both research design and service obligations requires considerable professional and scientific dexterity' (p.649). Whether the present study's transposition of operant principles was appropriate and justifiable is a question perhaps best answered by its outcome data.

(ii) The stance of radical behaviourism

The second of the themes identified by Blackman (1981) as central to the experimental analysis of behaviour is radical behaviourism. In this section we will examine what the system asserts in its approach to understanding behaviour.

At the centre of the radical behaviourist position is the assumption that 'behaviour can be understood in a real sense by relating/

relating it to the environmental circumstances in which it occurs, particularly to the differential consequences that patterns of behaviour may give rise to in specified conditions.' (Blackman, 1981, p.11). As a system of explanation, radical behaviourism is not confined to investigations of animal behaviour but is viewed as applicable to all aspects of human behaviour, including human interaction. Although behaviourism is frequently interpreted as a unidirectional account of behaviour, it is of the very essence of operant behaviour that it in turn affects the environment in which it occurs, for example by producing events which serve as reinforcers (Blackman, 1979). In any social situation, the behaviour of people in interaction is a subtle interweaving of influences, the one on the other. Such exchanges are open to operant analysis, although the continuous, reciprocal and interdependent relationships to be confronted here may present a different level of complexity from that encountered in the controlled laboratory setting.

The radical behaviourist seeks to understand human behaviour in terms of the functional significance of the pay-offs or consequences for what we do in particular settings. He would argue that such an account provides a useful explanation of behaviour, but one that may be too often overlooked. Such a position does not involve a denial of the existence of physiological or cognitive events but it does deny that such events are the special causes of behaviour. With regard to mental events Hallet (1981) interprets Skinner's aim as not to reduce or redescribe such things as thought processes in terms of behaviours. He views mental events as

explanatory/

explanatory fictions or inventions, and as unnecessary in the explanation of behaviour. They can be themselves conceptualized in terms of their possible environmental influences. For the radical behaviourist, therefore, understanding is sought through the functional relationships observed between behaviour and observable environmental events. Behaviour is thus viewed as valid subject-matter in its own right. Resorting to physiological or mentalistic explanations - while quite legitimate at another level of discourse - is seen as a deflection from a wholly justifiable attempt at a purely behavioural analysis (Boakes & Halliday, 1970). The radical behaviourist position is, of course, embedded in a view of science which deals only with publicly observable events and eschews mediational accounts of behaviour on the grounds that they are inferential. This is a brief exposition of the radical behaviourist stance. A fuller account is given by Skinner (1974).

From the point of view of treatment effectiveness, the philosophical underpinnings for behaviour modification are perhaps not crucial. It is important, however, to acknowledge that a philosophical paradigm can influence treatment practices. As Kuhn (1962) has suggested, the paradigm a scientist accepts can help to determine the kinds of research problems he finds important, the kinds of solutions he deems acceptable, and the ways in which he processes and evaluates evidence. Many criticisms have been offered of radical behaviourism and the kinds of interventions which it has generated - these will be examined later in the Discussion Chapter. Implicit in these criticisms are alternative models of understanding and intervention. Many of them call for important amendments to the radical behaviourist position. From a therapeutic perspective, the question of whether such amendments make for better outcomes is an empirical matter. Whether 'empirical' data/

data <u>can</u> settle such an issue remains to be seen, however, since in a field that is marked by rival theoretical stances towards therapeutic outcome, what is viewed as 'empirical' is often relative to respective theoretical positions.

The foregoing sections have outlined the main characteristics of the philosophical position underlying the operant approach. The next sections will deal with the principles and procedures themselves.

(B) The Principles of Operant Behaviour

(i) General considerations

The roots of most of the treatment procedures employed in the present study can be found in the operant approach to learning associated most closely with the work of B.F. Skinner (1938). A useful definition of learning is...'the process by which an activity originates or is changed through reacting to an encountered situation, provided that the characteristics of the change in activity cannot be explained on the basis of native response tendencies, maturation, or temporary states of the organism (e.g. fatigue, drugs, etc.)' (Hilgard & Bower, 1966, p.2) At the centre of Skinner's work are operants. He defined operant behaviour as that which an organism emits and which is affected by its environmental consequences (i.e. their future probability of occurrence is either increased or decreased by the events which follow them). This distinguishes operant behaviour from the more stereotyped and limited reflexes which are elicited by antecedent stimuli and which formed the basis of earlier experimental studies in Pavlovian or classical conditioning. Operant conditioning is the study of how such behaviour adjusts to different arrangements of environmental circumstances and consequences.

While the procedures employed in the present study were based on/

on operant rather than classical conditioning principles, in common with the bulk of classroom behaviour modification studies, it is clear that elements of classical conditioning principles are also of relevance to educational contexts. In the classical conditioning paradigm, neutral events come to elicit emotional reactions after being associated on numerous occasions with other events which produce those reactions. A teacher who reprimands or otherwise frequently punishes children may become a conditioned cue for fear and related negative emotional reactions in the children. In like manner, a teacher who reinforces children with positive events may come to produce desirable emotional responses in the children. Classical conditioning principles may be of no little importance in understanding children's emotional reactions in the classroom. Some of the terminology adopted within operant investigation makes the distinction between operant and classical conditioning rather confusing. Thus, examples of operant behaviour are usually called 'responses', but this is taken as defining the unit of behaviour studied, and it does not imply that the behaviour is bound to some eliciting stimulus as are the reflexes in classical conditioning (Blackman, 1981). Similarly, the environmental events, such as lights or auditory signals used in operant experiments, are referred to as 'stimuli', but they are not regarded as eliciting an automatic reaction as do the stimuli within classical conditioning experiments.

In the typical laboratory setting in which operant behaviour has been studied, animals such as rats or pigeons are exposed to environmental events which are consistently and carefully controlled by the experimenter. In a Skinner box, for example, extraneous and uncontrolled environmental variations are kept to a minimum.

A/

A limited number of events (e.g. lights, noise, food, water) are made available by the experimenter and the animal's behaviour is recorded in terms such as rate of key- or bar-pressing. It is within this kind of framework that the procedures employed in the present study were developed.

These procedures and the factors which affect their application will be outlined in the following sections. Fuller expositions are available in Reynolds (1975) and Honig & Staddon (1977). A variety of procedures was explored in the present study, as a means of achieving the principal objectives of (1) establishing or developing behaviour (e.g. volunteering information in class discussion), (2) accelerating or strengthening behaviour (e.g. sustained attention to class work) and (3) eliminating or weakening behaviour (e.g. disruptive noise in class). The methods of intervention and objectives will be dealt with fully in Chapter 6. The following sections will present the basic principles which helped to shape the classroom interventions.

(ii) Procedures for strengthening behaviour

A reinforcer is a stimulus which, when presented contingent upon the emission of an operant response, will tend to maintain or increase the probability of that response in the future. Reinforcement refers to the operations involved in such a procedure. A distinction is made between two kinds of reinforcing stimuli: positive reinforcers and negative reinforcers (Skinner, 1938). A positive reinforcer strengthens a response on being presented, while a negative reinforcer strengthens a response on being removed. Positive reinforcers are often equated with 'rewards' in common parlance but it is not so easy to grasp the concept of negative reinforcers/

reinforcers. To a degree, if such a stimulus has a reinforcing function when removed it must have aversive properties. Nagging, verbal reprimands and painful stimuli could function as negative reinforcers. However, whether an event functions as a reinforcer is entirely an empirical matter, and a function of the observed effects on behaviour. So despite the popular equation of reinforcement and reward, the impact of specific events on behaviour cannot be determined a priori nor should these events be assumed to have intrinsic hedonic qualities (Blackman, 1981). This is an important issue where behavioural procedures in applied settings are being considered. With children, for example, it is frequently assumed that sweets or toys will serve as positive reinforcers without any exploration of their functional relationships to the children's behaviour. Despite this proviso, and the need for a careful evaluation of the assumed properties of environmental events, it must also be acknowledged that events such as sweets, or an adult's attention, function as reinforcers for a great many children.

Reinforcers may also be categorised as either primary or conditioned. A primary reinforcer is one whose reinforcing properties do not derive from a history of prior conditioning. They are regarded as 'biological givens' and food, water and air are examples of potential primary reinforcing stimuli (Karoly, 1980). A neutral stimulus can be closely associated with a primary reinforcer and can eventually itself acquire reinforcing properties, when it is referred to as a conditioned reinforcer. It is often suggested, for example, that attention and affection become conditioned reinforcers for children through repeated pairing and association with/

with feeding processes.

The effectiveness of reinforcement operations is a function of many factors. These include the delay between a response and the delivering of reinforcement, the magnitude and quality of the reinforcer and the schedule of reinforcement.

Those responses which occur in close temporal proximity of reinforcement are learned more effectively than responses which are remote from reinforcement (Kimble, 1961). It is therefore recommended in most behavioural programmes that a reinforcer be provided immediately after the target response to maximise its effects. If reinforcement does not follow the response immediately, another intervening response may occur, and be unintentionally reinforced and strengthened. In relation to children's behaviour, this is an important consideration, in ensuring that the appropriate connections are made between the desired behaviour and the intended reinforcer. Delayed reinforcement, therefore, may contribute to imprecise and unfocused effects on behaviour.

The amount of reinforcement delivered for a response also determines the extent to which a response will be performed. Within certain limits, the greater the amount of a reinforcer delivered for a response, the more frequent the response (Kimble, 1961). However, if given in excessive amounts, a reinforcer may lose its effects. This is referred to as satiation, and is especially evident with primary reinforcers such as food, water and sex. Satiation may lead to a situation where these reinforcers may indeed become aversive but a period of deprivation usually restores reinforcing values. As Winkler (1971) points out, conditioned reinforcers such as praise and attention are also subject to satiation, but less so/

so than their primary counterparts.

While the amount of reinforcement is usually specifiable in physical terms, the quality of a reinforcer is a more elusive property. The quality of a reinforcer is determined by the preference of a person. Reinforcers that are more highly valued lead to greater performance.

Reinforcement schedules are the rules describing the manner in which consequences follow behaviour. In the simplest schedule, the response is reinforced each time it occurs. This is called continuous reinforcement, and is distinguished from intermittent reinforcement, in which only some instances of a response are reinforced. A behaviour is developed more effectively under conditions of continuous rather than intermittent reinforcement. However, once a piece of behaviour is well established, and an investigator is concerned to maintain it, an intermittent schedule makes it more resistant to extinction than a continuous one. In applied settings, therefore, reinforcement of every response is often recommended in the initial development of behaviour, changing to intermittent reinforcement once it is established. This is a basic distinction among schedule types, and numerous intricate experiments have been conducted to identify characteristic patterns of operant behaviour associated with other varieties of schedule (Williams, 1973; Zeiler & Harzem, 1979). Thus, high rates of responding are associated with ratio schedules, when the number of reinforcers delivered is expressed as a proportion of the number of responses emitted, whereas lower responding rates typically occur with interval schedules, in which reinforcement delivery is determined by the passage of time./

time. Additionally, rates of responding are more consistent over time when reinforcement occurs unpredictably in variable schedules than when it occurs after a specified number of responses or a specified period of time in a fixed schedule.

(iii) Procedures for weakening behaviour

The above reinforcement operations and characteristics of their application are of particular relevance in the development and strengthening of behaviour. We will now consider procedures which serve to weaken or eliminate behaviour, in situations, for example, where it might be regarded as inappropriate, excessive or dangerous. Operant behaviour may be weakened in three ways: by two forms of punishment procedure, or by extinction procedures (Bijou & Baer, 1978). Punishment is defined as an operation which serves to reduce the frequency of behaviour. It is therefore descriptive of an empirical relationship between an event and behaviour and does not necessarily entail pain or physical coercion as everyday usage of the term may suggest. In the first type of punishment, a negative reinforcer is made contingent upon the behaviour. This may take the form, for example, of a reprimand after engaging in some behaviour. Because of unpredictable side-effects (Azrin & Holz, 1966) punishment of this nature is little used in applied settings. In the second type, a positive reinforcer is lost or removed contingent upon the behaviour. This may take the form of time-out, where access to reinforcing events or activities (e.g. attention, company of peers, games, etc) is cut off for a limited period of time. A related procedure is response-cost, where previously held items of value such as money or points earned are lost contingent upon the response. The everyday equivalent would be a fine.

The/

The third procedure - extinction - involves no longer reinforcing a response that has been previously reinforced. It is equivalent to presenting neutral events where positive outcomes used to occur. A common example is the ignoring by a parent of the cries of a child which formerly elicited attention. Studies of the extinction process have shown that there are characteristic patterns of responding with an initial increase in the frequency, magnitude or intensity of a behaviour prior to its decline and eventual cessation (Ferster & Skinner, 1957). It is important to be aware of this possibility in treatment interventions, so that a parent or teacher employing 'ignoring' procedures is not trapped into further inappropriate attention when behaviour temporarily worsens.

(iv) Developing new and complex behaviour

The above procedures are basic ones in the strengthening and weakening of behaviour. A number of additional methods which are necessary for the development of new and complex behaviour need to be considered. These include shaping, chaining, prompting and fading.

In the development of new behaviour, reinforcement may often not be possible because the required behaviour does not occur. In shaping procedures, therefore,very small steps or approximations toward the goal behaviour are reinforced. These steps resemble the final response or include components of it. A shifting criterion for reinforcement is adopted so that closer and closer approximations are sought, while responses dissimilar to the goal are not reinforced. This is therefore an example of the combined use of positive reinforcement and extinction procedures.

Most/

Most human operant behaviour is more complex than simple stimulus - response units or associations. One way of approaching an understanding of complex behaviour is through the process of chaining. A chain refers to a sequence of responses in the repertoire of an individual which are bound together in a particular fashion. Solving a long division arithmetic problem, for example, involves the component processes of short division, subtraction and handling the remainder as a fraction. These processes would need to be taught separately and then linked together appropriately. A concept fundamental to understanding chains is that an event or stimulus which immediately precedes reinforcement becomes a cue or signal for reinforcing consequences. It is referred to as a discriminative stimulus. This sets the occasion for behaviour, but it also serves another important function, in eventually becoming a reinforcer itself, by virtue of its repeated pairing with reinforcement. In each link in the chain, therfore, apart from the first and the last, stimuli serve a double function of reinforcer and discriminative stimuli. Hence the links overlap, holding the chain together. The 'double function' serves as a reinforcement and as a cue for the next response in the chain. Chained responses, once learned and practised, operate smoothly and the partitioning of links is frequently difficult.

The concept of chaining is of central relevance to learning within an educational context. Gagné (1965) has specified five conditions essential for effective learning of chains. First, each stimulus-response unit, or link, must be learned separately before being/

being joined together. Second, each link must be executed in the proper sequence. Third, each link must be performed in 'close time' succession to ensure linking. Excesive delays may break the chain. Fourth, the sequence must be repeated sufficiently often until learning has been achieved and the chain becomes somewhat automatic. Finally, reinforcement must be present, since if the terminal link is extinguished, the chain breaks. The development of behaviour in the processes of shaping and chaining is facilitated by the use of cues, instructions, directions, models or examples in order to initiate a response. Events which help to initiate a response are called prompts. Prompts precede a response and when they initiate behaviours which are reinforced, they become discriminative stimuli for reinforcement. For example, if a teacher instructs a child to attend to his work, and his attention is appropriately reinforced, the instruction becomes a discriminative stimulus. The use of prompting plays an important role in shaping procedures, in ensuring that responses are initiated and allowing for more rapid approximations to the final response to be made. Although prompts may be necessary early on in the learning process, they can be withdrawn gradually or faded as learning progresses. Fading refers to the gradual removal of a prompt. To ensure that behaviour occurs without prompts it is necessary to fade and reinforce the responses without cues or signals.

(v) The role of antecedent events.

Operant behaviour was defined at the outset of this discussion as behaviour which is influenced by its consequences. However, it is apparent from consideration of the nature of discriminative stimuli/

stimuli that antecedent events also influence behaviour. Prompts represent a group of events which precede and facilitate response performance. In addition, reinforcement may occur in some situations or in the presence of certain stimuli, but not in others, in a process of differential reinforcement. Thus, certain settings may signal that a response is likely to be reinforced, while others will signal that reinforcement is unlikely. When the individual responds differently in the presence of different stimuli, he has made a discrimination. When responses are differentially controlled by antecedent stimuli, behaviour is said to be under stimulus control.

It is often observed in school settings that children will studiously observe the instructions of certain teachers but blithely ignore the same instructions from others. They can be seen as having formed a discrimination and as responding to the differently perceived consequences for their actions. In interventions in applied settings, treatment may be carried out in certain settings and not in others, e.g. in the school, and not in the home, or in one classroom but not others. Where discriminations are formed as to the settings in which reinforcement is likely, and behaviour comes under the stimulus control of treatment settings, it may be found that desired changes occur only in those settings in which reinforcement has been available. This problem introduces a concept that is of fundamental importance to treatment interventions - generalisation of learning. It is one thing to bring about change within the treatment situation, but it is another matter to ensure that learning transfers to other situations. Of related concern is the influence of change in one learned response on other responses.

If/

If a response reinforced in one setting also increases in other settings, even though it is not reinforced in these other settings, this is referred to as stimulus generalisation. Stimulus generalisation can be regarded as the opposite of discrimination. Where discrimination occurs, the response will not generalise across situations. The degree of stimulus generalisation is a function of the similarity of the new stimulus, or situation, to the stimulus or situation under which the response was trained. Response generalisation refers to learning in which alteration of one response influences other responses. For example, if a child's aggressive hitting is reduced, it may be found that shouting and swearing also diminish. Skinner (1953) suggests that the reinforcement of a response increases the probability of occurrence of other responses which are similar.

As Kazdin (1977) points out, these generalisation concepts are more than descriptive terms, and are often taken to imply an explanation for the spread of effects. However, although it might be stated that stimulus generalisation is a function of the similarity of conditions across which the response transfers, there is rarely any demonstration that such similarity exists.

This completes the summary of the principles underlying the intervention procedures which will be described in Chapter 6 . In the move from the laboratory, where these principles have been elaborated, into less controlled settings, as in the present study, some difficulties in translation may arise. The next section will examine some of these issues.

(vi) Moving from the laboratory to less controlled settings.

The foregoing discussion of learning principles suggests a degree of precision and detail in application which may well be readily attained in the laboratory, but which may prove unrealistic in the less controlled natural setting of the classroom. The introductory chapter indicated some of the differences that might exist between 'demonstration' or special research projects, and natural settings, and the need to test out behavioural approaches in these less controlled settings. The theme of the present section is that such a move may well involve compromises and adjustments at the level of basic operations.

It was recognised in the present study that pure translations of procedures, and their rules of application, would not be possible. This was evident in a number of ways. For example, the most effective application of positive reinforcement by a teacher might call for reinforcement to be provided immediately once the desired response has occurred. Given that the teacher could have 30 children in her charge, and be immersed in other pressing activities, the ideal may not often be realised. That example reflects the practical difficulties in controlling the natural environment of the classroom. In such a complex social setting, there may be many influences on behaviour and it is difficult to isolate one from another. Thus, 'extinction', as a method of withholding reinforcement for, say, mildly disruptive behaviour, typically involves withholding teacher attention. Extinction is often identified with such a strategy in classroom studies, yet 'true' extinction, which is readily achieved in the laboratory, may not be occurring at all in the classroom/

room , where other sources of reinforcement (e.g. peer attention) may be competing or dominant. Both these examples of translation difficulty suggest that the procedures as described may be only approximations which fall short of ideal requirements. Stringent experimental analysis may also prove difficult. Thus, it was also recognised that the demands of dealing with complex behavioural problems might call for a 'package' programme involving several procedures, rather than an individual technique, so that identification of the effective element is complicated. Such difficulties in the translation of laboratory concepts into 'real-world' application have been commented on by other writers. Azrin (1977),for example, condemns the naive view that 'one need only translate the basic reinforcement concepts into clinical practices and a "cure" automatically appears'. (p.143) Describing the development of treatment procedures within his own clinical practice, he suggests that 'For all of the treatments, changes in the initial reinforcement conception were required, and although the additional procedures were often derived in turn from a reinforcement framework, their necessity was not predicted by the model. The final treatment program in every case required improvisations, detours, and innovations because of problems unanticipated by the reinforcement analysis.' (p.143) The basic laboratory-derived principles should therefore perhaps be viewed as providing an organising framework for an intervention rather than one which can hope to totally define it. Macmillan & Morrison (1980) suggest that in early behavioural applications within special education, there were efforts to 'force' practices/

practices into constructs and principles of learning, because of the insecurity associated with 'less-than-pure' applications. The more recent willingness to describe what is actually taking place, rather than employing a construct that differs subtly, but significantly, from what is actually happening, is seen as a sign of maturity.

(vii) Concluding Note

The present chapter has outlined the theoretical basis for the intervention procedures adopted in the present study. The philosophical stance of the behavioural approach has been described, along with the operant principles and procedures which it has generated. The emphases on study of individuals and the pursuit of thorough-going behavioural explanations can be seen as strengths of the approach. The present research can be viewed as drawing on these strengths, but the requirements of intervention within applied settings, with large numbers of subjects, may call for a readiness to operate in conditions which diverge somewhat from the theoretical and procedural optimum.

CHAPTER 3

THE EVALUATION OF THERAPEUTIC OUTCOME

THE EVALUATION OF THERAPEUTIC OUTCOME

Introduction

The focus of the present study is behavioural intervention with children experiencing adjustment problems within an educational context. However such interventions in educational settings need to be viewed against the wider background of treatment research with children, and in the context of the methodological issues which are relevant to child therapy research in general. In the present chapter, the first of these requirements will be approached by consideration of (a) studies of child psychotherapy, (b) behavioural interventions in other than educational settings, (c) comparative studies where behavioural and other treatment modalities are explored and (d) meta-analysis, as an innovative approach to evaluation. This survey will be necessarily selective. It will draw on the available reviews and emphasise experimental investigations rather than the numerous case studies of single-case reports which abound in the field. (This is not to suggest that such studies are irrelevant or scientifically worthless. As Kazdin (1981b) has pointed out, case studies and experiments can be viewed as falling on a continuum that reflects the degree to which scientifically adequate inferences may be drawn.) Only studies concerned with emotionally disturbed or maladjusted children are included, while investigations of mentally handicapped, delinquent and psychotic children fall beyond its scope.

A question of major interest in this survey is whether there is/

is sound evidence available which can serve as a standard of therapy effectiveness against which the success of the intervention in the present study can be judged. The review of treatment studies will be followed by a more detailed examination of the methodological issues which are relevant to the present treatment investigation. The main literature review dealing with treatment interventions in educational settings will be presented in Chapter 4.

A. Investigations of Therapy Effectiveness

(i) Child psychotherapy studies.

The term 'psychotherapy' is often used generically to refer to the wide range of treatment approaches that are available within the helping professions. In this sense, 'psychotherapy' can be regarded as subsuming behaviour modification methods, but the present section is not concerned with treatment approaches which are explicitly based on learning theory. The task of defining psychotherapy in an adequate fashion is no easy one. A frequently cited definition is that offered by Meltzoff and Kornreich (1970):

> 'Psychotherapy is taken to mean the informed and planful application of techniques derived from established psychological principles, by persons qualified through training and experience to understand these principles and to apply these techniques with the intention of assisting individuals to modify such personal characteristics as feelings, values, attitudes, and behaviours which are judged by the therapist to be maladaptive or maladjustive' (p.4).

Such a definition may be regarded as generally applicable to child as well as to adult psychotherapy. It is important, however to/

to recognise the diversity of approaches that might lie behind any proffered definition. 'Psychotherapy' is not a unitary procedure, and in clinical practice, it is likely that several approaches may be combined. Thus, a child in individual play therapy may also be given medication, he might be involved in family therapy, or the parents might be counselled separately. While such variability would be more tightly controlled in research studies than in routine clinical work, it creates a degree of ambiguity in interpreting reviews of early studies of child psychotherapy.

The most influential reviews of the effectiveness of psychotherapy with children have been those of Levitt (1957, 1963; Levitt et al., 1959). In his first review (Levitt, 1957) 18 studies were examined involving some 8000 children. The ages covered by these studies ranged from preschool to 21 years but with very few cases over 18 (median age 10 years). Most of the cases were classified as neurotic. Two-thirds were rated as improved at end of treatment and slightly over three-quarters at follow up. However, Levitt also presented data suggesting that the base rate for improvement in untreated children was essentially similar and concluded that his review lent no support to beliefs in the usefulness of psychotherapy. This conclusion requires some examination. His untreated comparison groups were derived from two studies: Witmer and Keller (1942) and Lehrman et al., (1949) and consisted of 'defectors' from treatment, i.e. patients for whom treatment was recommended but not accepted or/

or completed. The question is immediately raised of the equivalence of treated and untreated cases and the possibility that defectors constituted a biased control group. Levitt presented evidence that the defectors and treated cases did not differ on some 61 factors, 'including two clinical estimates of severity of symptoms' and eight other factors relating to symptoms (Levitt, 1957). This procedure is viewed with some scepticism by Hartmann et al., (1977) who point out that failure to reject the hypothesis of no difference is not equivalent to asserting that no differences exist. Witmer and Keller indeed acknowledge that half the cases came from either the courts or physicians with a request for a diagnostic service only, suggesting that the children may not have been seriously disturbed. Furthermore, it is also suggested that some of the children may have received treatment. A further criticism of this study is that the follow-up period of 8 to 13 years could have covered many fluctuations in psychological well-being, and it is indeed questionable whether such a lengthy follow-up with young children can provide sensitive and meaningful data.

Close examination of the Lehrman et al., (1949) study also casts doubt on the claim that it can make a sound contribution to establishing a reliable base rate. Thus, the untreated children are found to have been exposed to diagnostic interviews which may well constitute therapeutic influence (Miller et al., 1972). They were reported to have more adequate parents and the rating procedure employed is not adequately described. It is/

is possible, therefore, that the 'baseline' improvement rate of 72.5% derived from these studies is inflated, or at the least, unreliable. It was, however, in keeping with the tenor of the times, following Eysenck's (1952) claim of a two-thirds improvement rate in untreated adult neurotics, similar to that in treated cases. It was also Levitt's conclusion following a further analysis of the outcome of some 1000 child guidance cases, again comparing treated children and defector controls (Levitt et al., 1959). This is in some ways a curious study to test the effectiveness of psychotherapy with children, as Barrett et al., (1978) point out, in that the child alone was the focus of treatment in only one-tenth of the cases, and the mother was the only family member treated in more than 40% of all cases. In other words, in nearly half the cases involved, the child was not directly treated. Levitt himself comments on the number of different kinds of therapists involved in these studies, with psychiatrists, social workers and teams of clinicians operating at different points in the patients' milieu. The additional diversity in treatment approaches underlines the difficulty in interpreting the results of surveys of studies.

A further review of 24 studies confirmed for Levitt (1963) that two out of three children are improved at the end of treatment. This review examined outcome within diagnostic groups and concluded that outcome was best for neurotic cases and worst for children with acting-out symptoms, for whom a 55% improvement was cited. The overall two-thirds improvement rate for neurotic children is possibly inflated by the inclusion of many cases with rapidly improving problems, such as school-phobia and enuresis.

Although Levitt's conclusions as to the ineffectiveness of child psychotherapy may need some qualification, there is little substantial evidence to the contrary. Eisenberg et al., (1965) in a controlled study of out-patient treatment, demonstrated some beneficial effects of brief psychotherapy, but rating bias could not be ruled out, and the treatment effects, though statistically significant, were small. Two studies which sidestep the issue of baseline improvement rates - Rosenthal and Levine (1971) and Heinicke (1969) - have presented conflicting findings on the issue of psychotherapy 'dosage' - frequency of sessions and duration of treatment. The first reports no better results for long-term as opposed to short-term psychotherapy and is weakened by problems of attrition. The second claims a superior outcome for more frequent treatment (once weekly versus four times per week) although the outcome measures do not consistently favour one group over the other. There were also only four cases in each group.

A recent review of psychotherapy outcome studies with adolescents (Tramontana, 1980) presents a more optimistic picture than do the reviews of Levitt. Thirty-three studies were examined indicating a median positive outcome for about 75% of treated cases, compared with a rate of 39% without psychotherapy. These figures, however, have to be treated with caution. As the author notes, many of the studies are of poor methodological quality. For example, treatment other than psychotherapy was also provided in a number of the investigations reported so that interventions are/

are confounded. In seven studies, the nature of the psychotherapeutic intervention was unspecified, creating uncertainty as to what was being evaluated. The criteria for inclusion in the review also seem somewhat unsatisfactory since in several of the studies, 'psychotherapy' took the form of hypnotherapy or behavioural approaches. The low improvement rate in untreated controls is of interest, given its marked divergence from the rates we examined earlier. The generality of this rate across the span of adolescent problems is in doubt, however, given the methodological inadequacies of many of the studies and the fact that over half the controlled investigations involved delinquent cases. The outcome for such cases may be particularly poor (Robins, 1979).

It was noted earlier, in discussion of Levitt's (1957) first review, that somewhat better rates of improvement were observed at follow-up than at termination (three-quarters of cases as against two-thirds). Wright et al., (1976) have gone on to argue that a very different impression of the effectiveness of individual therapy is obtained if one considers results at follow-up rather than results at close. Reviewing six studies employing experimental controls (Lehrman et al., 1949; Seeman et al., 1964; Heinicke, 1969; Love et al., 1972; Dorfman, 1958; Miller et al., 1972) only one study out of the six provided evidence in favour of psychotherapy at the close of treatment. However, at follow-up, four of the six studies indicated improvement, one showed no change, and one showed deterioration. The authors suggest that/

that improvements at follow-up may be most common when psychotherapy sessions number thirty or more, lending support to the findings of Heinicke's (1969) study, cited earlier.

The authors acknowledge the presence of methodological problems in their sample of studies, the most common being potential bias in improvement ratings, but their review nevertheless strongly underlines the need for follow-up measures in therapy research. As to possible explanations of such increments at follow-up, Wright et al., propose that such findings are most consistent with theories which view psychotherapy as affecting central, underlying, or structural aspects of personality functioning, rather than directly altering behaviours. They do not consider the equally plausible explanation, however, which implicates behaviour change rather than change in central processes. Thus, it is possible that where behaviour change is achieved in therapy, further improvement can be viewed as a cumulative process, as changed behaviours elicit favourable responses from others. Continued progress may therefore occur in a cycle of reciprocal positive feedback.

What may be concluded from this part of the review? The clearest conclusion seems to be that there is no firm evidence that psychotherapy is effective. However, there are some suggestions that rates of improvement may be better in the longer term (at follow-up) than at treatment termination, and follow-up evaluations are strongly recommended. Also lacking in the field/

field is agreement as to the rate of change that may occur in the absence of treatment, which could act as a reliable standard of comparison for treatment studies. With regard to baseline improvement rates without treatment, it is likely that the traditionally accepted figure of two-thirds improved may need to be revised downwards, although perhaps not as far as Tramontana's review for adolescents would suggest. It is of interest that a drastic revision of the two-thirds improvement rate for adults has been argued by Bergin and Lambert (1978). They claimed a median improvement rate of only 30% from their own review but the original position has been vigorously restated recently by Rachman and Wilson (1980).

(ii) Behaviour modification studies.

This section will focus on studies of behaviour modification conducted outside of educational settings. Of the research that has been conducted within the behaviour modification model, the work of Gerald Patterson and his associates has been particularly influential. A detailed description of their approach to family intervention has been provided by Patterson, Cobb and Ray (1973). Patterson (1974) reported on interventions with twenty-seven children, referred by community agencies for problems subsumed under the label of 'conduct disorder'. Treatment was conducted in the home with the parents as mediators and follow-up data were collected at periods of up to 12 months. Fourteen children also received treatment in the classroom, via teaching staff. Patterson/

Patterson concluded that the interventions 'were moderately successful in producing reductions in noxious behaviour' and that follow-up data 'showed that these improvement rates persisted' (p.479). This claim has been disputed by Kent (1976) who argued that the sixteen children who were available at the 12-month follow-up comprised a biased sample. The eleven who dropped out showed 2.39 times more deviant behaviour on observational assessment than the remainder of the sample before treatment. In a rejoinder to this critique, Reid and Patterson (1976) demonstrate that for those subjects who completed the follow-up but showed initially high levels of deviant behaviour, a statistically significant persistence effect is obtained. However unequivocal interpretation of this effect is still difficult, since no control group was included in the study.

Although two other studies following this intervention model (Walter and Gilmore, 1973; Wiltz and Patterson, 1974) have provided control group data and claimed significant success for the treated group, in neither study were treatment and control group data directly compared. Instead, it is reported that deviant behaviour in the treatment group decreased while in the control group it did not. Baseline data in both studies show that controls displayed higher rates of deviant behaviour, further complicating interpretation of outcome. Unlike Patterson's (1974) study these investigations also provided data for only 4-5 weeks. In a further exploration of the Patterson model, Ferber et al., (1974) found that only three out of seven families of unduly aggressive/

aggressive children showed positive short-term improvement, with only one family showing substantial change at a one-year follow-up. More recently, Fleischmann (1981) has reported a successful replication, with significant improvement at both termination and one-year follow-up. As with the original study, no untreated controls were available.

The above studies focused specifically on children with conduct problems. O'Leary et al., (1973) have reported on out-patient intervention for seventy children with a wide range of presenting problems, including social withdrawal, temper tantrums, delinquent acts, immaturity, phobias and conflict in family relationships. Behavioural treatment was available directly with some children, or through contingency management procedures with parents as mediators with others. At the end of treatment (median duration 3½ months) 87% of cases were rated as improved by the therapists and 90% by parents at follow-up, on average, six months later. While the authors acknowledge the absence of control data, they point to the markedly higher improvement rate than that reported by Levitt (1963). However, it is unclear what clinical significance can be attached to the changes reported and the source of the improvement ratings raises questions of bias.

A number of studies are available which suggest support for behavioural interventions with fear- and anxiety-related problems, e.g. fear of the dark (Leitenberg and Callahan, 1973), fears of dogs (Obler and Terwillinger, 1970; Bandura et al., 1967)/

1967) and of snakes (Ritter, 1968). The clinical significance of the problems treated in these studies may be open to question. Children identified by teachers as socially withdrawn and isolated were successfully helped by the use of filmed modelling techniques in a controlled investigation by O'Connor (1969). However the beneficial effects of this procedure were found not to be maintained at follow-up (Keller and Carlson, 1974), and Gottman (1977) completely failed to replicate O'Connor's findings.

As with the psychotherapy studies examined earlier, methodological questions complicate interpretation of these reports of behavioural interventions, so that no unequivocal picture emerges. It should be noted that the vast body of behavioural research, conducted in within-subject designs, is not considered here. It is also apparent that, with the exception of the O'Leary et al., (1973) report, improvement is evaluated in terms which offer no possibility of direct comparison with psychotherapy studies. This issue of comparative effectiveness is an important one, and central to the present study. We now turn to an examination of these studies - unfortunately few in number - which have addressed this.

(iii) Treatment comparison studies.

As a method of evaluating therapy effectiveness, designs which contrast a single form of treatment with no-treatment are capable of producing only limited conclusions, since a successful outcome may be attributed to the fact of having intervened in any way at all, rather than to the specific treatment procedures employed/

employed. A somewhat more refined approach compares different treatments and goes some way towards identifying specific ingredients which may be of help. Comparative studies of this nature with children are rare.

Humphrey (1966) compared the effectiveness of psychotherapy (N = 17) and behaviour therapy (largely desensitisation) (N = 20) with clinic cases. The cases were randomly allocated to the two treatment groups, and a control group of 34 children was drawn from defectors from treatment. The measures of treatment effectiveness included therapist ratings, blind evaluations by a consultant psychiatrist, and teacher ratings. Both forms of treatment produced a better outcome than the controls (of whom 66% improved) and behaviour therapy was marginally superior to psychotherapy both at end of treatment and at 9-month follow-up. The improvements in the behaviour therapy group were also achieved more economically (mean of 17 as against a mean of 31 psychotherapy sessions). A number of design problems limit confidence in these results: all treatments were conducted by the same therapist, creating the possibility of bias; it is not clear that the treatment modalities differed sufficiently in their application; the treatment groups differed in pre-treatment ratings of severity, and, finally, consistent differences were found in the ratings of the independent assessors. The study is nevertheless an ambitious one which has, unfortunately, not been replicated.

As a treatment comparison study, the above investigation is similar in design to the intervention which is the focus of this/

this thesis, as is the work of Miller et al., (1972). This study examined the impact of psychotherapy and desensitisation on 44 phobic children, 6 - 15 years old, on the basis of time-limited treatment (8 weeks at 3 sessions per week). There were 23 waiting list controls. The two treatments did not differ in effectiveness but both differed from the results of the controls, on both independent clinical ratings and behaviour ratings by the parents. At a two year follow-up (Hampe et al., 1973) the rate of success for younger children was considerably better than that for older children, and some 73% of the treated groups were considered successful as opposed to 34% of the untreated group. While this study concentrated on a specific childhood disorder - phobic reactions - the divergent outcomes in treated and untreated children is intriguing, particularly the limited progress of the latter. In this respect it stands in marked contrast to the earlier studies and reviews we examined.

While this is a commendable attempt to make controlled treatment comparisons, a number of problems make interpretation of the findings difficult. First, the severity of the phobias is questionable. Over 43% of the school-phobic children were attending school, and school-phobics comprised some 69% of the sample. Second, treatment procedures may have overlapped excessively (e.g. behavioural strategies were employed in the psychotherapy group) so that clear enough distinctions may not have been possible in the evaluation. Third, rating procedures and outcome test results are not described in detail.

Neither/

Neither of these treatment comparison studies, therefore, points to the clear superiority of one form of treatment over another. Both, however, suggest that treatment is preferable to none. The paucity of studies in this category reflects the lack of attention given to comparative treatment effectiveness with children.

(iv) Meta-analysis: an innovative approach to evaluating therapy outcome.

A cursory survey of the treatment outcome literature is sufficient to convince the reader that a veritable minefield of conceptual and methodological problems confronts the therapy researcher. It is difficult enough to draw firm conclusions from single studies, even more so when an overall review of such studies is attempted, and an accumulation of studies of variable quality are lumped together. Such surveys run the risk of over-estimating the effectiveness of therapy when one considers that publication policies may contribute to an over-representation of successful studies in the literature (Tramontana, 1980).

An approach to therapy evaluation which seems to compound the above problems is the recent meta-analysis method of Smith and Glass (1977). This was an attempt to make overall statements on therapy effectiveness and the comparative effectiveness of different therapeutic approaches by means of a thoroughly comprehensive review of available studies. Although studies involving children were included in the analysis, findings for these cannot unfortunately be addressed directly since they were included in/

in the global conclusions and not dealt with as a sub-group. The basic unit of this analysis is an 'effect size' defined as the mean difference between the treated and control subjects divided by the standard deviation of the control group. More than 830 effect sizes were computed from 375 studies. Smith and Glass concluded that the average study showed a two-thirds standard deviation superiority of the treated group over the control group. Thus the average client was 'better off than 75% of the untreated controls' (p.10) This was interpreted as reflecting the effectiveness of therapy. On examining studies in which behavioural and non-behavioural approaches were directly compared, it was concluded that there were no essential differences in effectiveness.

This report has been the target of heated debate eliciting support (Shapiro, 1980) but also fierce criticism (Eysenck, 1978; Rachman and Wilson, 1980). A fundamental problem with such an analysis is that the findings can only reflect the quality of the original data: where the original studies are inadequate, a meta-analysis which lumps together widely varying investigations may only add to the confusion. Amongst the criticisms levelled by Rachman and Wilson (1980) is that many relevant comparative outcome studies - particularly behavioural ones - have been omitted from the Smith and Glass analysis while reliance is placed on others of highly dubious quality and relevance, or available only in unpublished dissertation form. Studies of widely varying quality are, in fact, given equal weighting. Furthermore, a number/

number of 'uniformity myths' (Kiesler, 1966) are subscribed to. Firstly, in relation to measures of therapy outcome, the whole gamut of outcome measures is mixed together on the basis that they are more or less related to 'wellbeing' and at a general level are therefore comparable. Secondly, in relation to treatments, the necessary discriminations within umbrella terms such as 'behaviour modification' and 'eclectic therapy' are not made.

Psychotherapy - in its different forms - may well be effective, and some forms may be more helpful than others, but it is doubtful that valid and widely-accepted judgements on these possibilities can be drawn from meta-analysis. It would seem preferable to base such judgements on the results of well-controlled studies. It is evident that methodological difficulties have dogged many of the investigations considered in this section. It is appropriate that some of these issues should now be addressed in more detail.

B. Methodological Issues.

There are numerous concerns related to method and design in therapy outcome research and only the issues of direct relevance to the present study will be dealt with here. Comprehensive reviews are available in Hartmann et al., (1977), Bergin and Lambert (1978), and Gottman and Markman (1978).

(i) Spontaneous remission.

We have already noted the emphasis placed on the assessment of baseline changes without treatment, or, as the phenomenon is/

is more widely known, 'spontaneous remission'. The problems in this area have generated much heated debate within psychotherapy research, perhaps more with reference to adults than children.

One may with some justification question the usefulness of studies of spontaneous remission and the meaningfulness of the data they generate. Although the term strictly refers to change in the absence of any formal therapeutic intervention (Rachman and Wilson, 1980) rather than change occurring without cause, it reflects a rather blinkered conception of how therapeutic change may come about. Most of what we know about change is within the context of formal therapy yet it is abundantly clear that many people in search of psychological help do not choose mental health professionals but obtain help from other sources (Gurin et al., 1960). Some 50% of clients drop out of treatment on an unplanned basis after the first interview and up to 80% before six interviews (Fisher, 1978), and it is likely that many of these will seek help elsewhere. Yet we are relatively ignorant about the kinds of factors outside of therapy which may be of benefit and they tend to be the object of speculation rather than research. Spontaneous remission rates may therefore obscure more than they reveal. More specific criticisms of spontaneous remission studies can be focused on decisions as to which studies or patient samples are relevant, the criteria employed for classifying improvment, and questions of the reliability/

reliability of diagnostic procedures employed. The whole practice of lumping together studies of different patient groups, with variable follow-up time spans, and representing remission with a single figure may be more confusing than helpful. These are important considerations in evaluating therapy outcome research but specific problems arise where children are concerned. With children, developmental change and variability are unavoidable processes which have to be separated from therapeutic effects. Levitt (1971) has pointed out that many children's problems are essentially developmental and disappear without organised intervention. Many 'normal' children show behaviours which are usually considered indicative of emotional disturbance - temper tantrums, sleep disturbance, enuresis, hyperactivity, specific fears, etc (Macfarlane et al., 1954; Lapouse and Monk, 1959). Levitt suggests that most children 'grow out' of these difficulties so that 'treatment', if applied, may be unnecessary. Referral for help may be a function more of the parents' reaction than of the child's need for treatment (Shepherd et al., 1966). Given the natural trend to improvement, the availability of untreated controls is an important consideration in therapy research with children.

A further developmental issue is that 'symptom substitution' may occur so that the child may 'grow out' of one problem but develop another at a later stage. The time span over which change is monitored is therefore quite critical, if sensitive measurement of changes are to be made. Furthermore, measurement at several time points would seem preferable to single probes after treatment. The/

The selection of assessment procedures should also cater for the possibility of symptom shift, perhaps by including multiple measures of change, otherwise conclusions on improvement based on measurement of the initial problems may be quite erroneous. A final important consideration relevant to these changes is the age of the children in such studies, since particular problems are characteristic of certain ages and not others.

(ii) Experimental design.

The main task facing the investigator in the area of experimental design in therapy research is to develop appropriate control strategies which will enable him to rule out plausible rival hypotheses in the attempt to identify 'true' treatment effects. The range of designs available afford differing degrees of control (see Campbell and Stanley, 1966; and Kiesler, 1971 for comprehensive coverage). Hartmann et al., (1977) classify the available designs as suitable for either exploratory or confirmatory purposes. In the former category belong traditional case studies, descriptive investigations, quasi-experimental designs, and other naturalistic correlational designs. Typically, these exploratory designs do not provide the degree of control for rival hypotheses required for a comparative treatment study, such as the present one.

Two general classes of confirmatory designs are available. One class consists of individual-subject designs, the most common forms being variants of the reversal or ABAB design and the multiple-baseline design (Gelfand and Hartmann, 1975). The second class consists of a variety of true experimental group designs. This/

This includes a range of designs involving random assignment of subjects to conditions, and is exemplified by the treatment versus no-treatment design, the comparative-treatment design, and designs employing various numbers and combinations of treatment and control groups.

Individual-subject designs have played a fundamental role in behaviour modification research, and have considerable strengths where functional relationships are sought between manipulated and dependent relationships. However, as discussed in Chapter 2 these designs were not considered appropriate to the needs of the present study. They were ruled out because of the requirements of large group comparisons, long-term follow-up and the difficulties of introducing reversal procedures in regular educational settings.

A major concern in employing a confirmatory group design has to do with the nature of the control group. Earlier sections in this chapter, dealing with issues of spontaneous remission and natural developmental changes in children, point to the need for a carefully selected untreated control group. Difficulties have already been noted with naturally defined control groups such as treatment defectors or terminators. Random assignment to treatment and control groups is the preferred option (Hartmann et al., 1977). This indeed was the procedure adopted in the present study, with random allocation to three treatment groups, and an untreated control group. Random allocation, however, does/

does not guarantee that subjects are initially equivalent. Where groups are not evenly matched, interpretation of treatment results is complicated. Take for example, an experiment in which the treated group improved more but also initially demonstrated more disturbed behaviour than the control group. It cannot be convincingly argued here that treatment was effective since competing hypotheses based upon regression to the mean and differential maturation are equally viable. The statistical phenomenon of regression to the mean is a most important matter in therapy research. If a group of cases is selected by an extreme score on any assessment procedure, that group will inevitably give a less extreme score if the assessment is repeated - the scores regress to the mean of the population. If two groups are being compared, and their initial scores are not the same, then differential regression may account for observed differences on post-testing. The procedure of analysis of covariance allows adjustment to be made in final scores which will allow for differences in initial scores. These procedures of statistical control were employed in the present study.

Other difficulties may be observed with untreated control groups. First, allocation to such a group does not ensure that help is not received. In addition to unprogrammed help, which may be obtained elsewhere, the support and reassurance that may be provided in the process of intake and follow-up may serve to reduce differences in outcome between treated and untreated subjects (Meltzoff and Kornreich, 1970). Second, withholding of/

of treatment raises ethical questions. However, such action does seem justifiable if evidence for the treatment's efficacy is lacking, and given that some forms of treatment may indeed not only be unhelpful but worsen a client's adjustment (Bergin and Lambert, 1978). Third, problems of contamination may arise when control subjects have contact with treated cases, so that the possibility of transmission of treatment effects is present. There is particular risk of this occurring where control and treatment programmes are run in proximity within settings such as hospitals or schools where not only may subjects have contact but also the staff involved in treatment implementation may be in contact with control subjects. This problem is of particular relevance to the present study, in that both treatment and control groups were located in the same schools. However, it is to be noted that transmission of beneficial effects to the control group would lessen the chances of detecting 'true' effects in favour of the treated group.

A fourth difficulty which may be encountered with untreated control groups involves the control client's perception of his status. When clients are placed in a control group, the understanding of the allocation that they develop may serve to complicate their 'control' status. Wortmann (1975), for example, suggests that controls may believe they were placed in the control condition because they have superior prognostic characteristics and may thus work harder at resolving their problems - a reaction which could obscure real treatment effects. Alternatively, they may be embittered by not receiving treatment, which may worsen their/

their adjustment. Such a response would make an ineffective treatment seem effective. In practice, it is difficult to establish to what extent these biases may be operating, and seeking such information may require an undesirable degree of contact with the controls. In the present study, very few if any of the control cases would have had any knowledge of their formal 'allocation', so it is unlikely that these difficulties operated to any significant degree. The issue of clients' expectations raises the question of how these might be adequately dealt with, and the use of attention-placebo control groups has been suggested (Kazdin and Wilcoxon, 1976). Such groups may also be used to control for factors such as frequency of contacts, and therapeutic interest. Hartmann et al., (1977) suggest that the use of these 'active' control group conditions may be limited to the investigation of brief psychotherapy, since the long-term use of attention-placebo controls taxes both the creativity of clinical researchers and the credulity of participants. It is also evident that such controls may tax manpower resources, which may already be stretched to cover treatment and assessment requirements.

(iii) The assessment of change

The measurement of change in therapy provides an area of debate which gives full rein to differing theoretical conceptions of what processes are involved and how they may best be tapped. Investigators from different perspectives seldom agree on what constitutes a meaningful outcome. Measures of change tend to be theory-specific and not uniformly accepted across different models of therapeutic intervention. There is considerable agreement/

agreement, however, that global and unidimensional judgements by therapists that clients are 'improved', 'partially improved' or 'unimproved' will not produce meaningful information. Most writers in the field also now stress the need for multiple measures of change which reflect the complexity and diversity of the processes which may be involved. Johnson and Eyberg (1975) , for example, suggest that 'no one class of assessment instruments can adequately reflect change in complex human processes. All types are fallible and worthy in different respects, and each may be best employed to complement, clarify and correct for the others.' (p. 918). Franks and Wilson (1980) indicate a further important reason for being concerned with multiple-method assessment. They point out that 'data can be more influenced by the method used to collect them than by variations in the topic being assessed. Unless measured via different procedures, one cannot isolate and remove from the total variability that which is due to method variance.' (p. 81) Mintz et al., (1973) have also commented on sources of random error in such measurements. Attention to the reliability of measurement instruments is an obvious necessity.

The commonly observed discrepancies between different outcome measures (see Chapter 1) may reflect the different domains of functioning which might be tapped (e.g. behavioural, physiological, cognitive), varying settings where measures are conducted (e.g. school, home, clinic) and different times at which they are obtained. Inconsistencies/

Inconsistencies between outcome measures are therefore to be expected and it is only unwarranted assumptions concerning homogeneity of change processes which have led researchers to expect otherwise. Such beliefs would tend to a view of low correlations between measures as reflecting the low reliability of the measures, and encourage their rejection in favour of those which provide a more reassuring degree of convergence.

Discrepancies between different measures may also reflect the differing vantage points of those who evaluate change and the values of those who occupy each vantage point (Strupp and Hadley, 1977). They suggest that divergences in viewpoint and values are inevitable but nonetheless legitimate. They identify three major 'interested parties', who may be concerned in such evaluations: (1) society, which includes significant persons in the person's life. Where children are concerned, these may include parents and teachers. (2) the individual patient and (3) the mental health professional. No one of these can be regarded as having absolute priority in its viewpoint. Society's agents may be concerned primarily with the maintenance of an orderly world and may stress criteria such as stability, predictability and conformity. These may be at odds with the individual's criteria which will focus on a subjective sense of wellbeing. This component of Strupp and Hadley's model cannot be applied in a straightforward fashion to children who may often be brought for treatment by parents, or referred for intervention by teachers though firmly asserting themselves that they are not subjectively distressed. It is important to provide some opportunity for self-report, although sensitive and adequate accounts of subjective feelings/

feelings may be difficult to obtain. From their vantage point, mental health professionals will bring varying conceptions of human functioning to bear on the issue of evaluation. Therapists' ratings of improvement have been probably the most frequently employed measures in therapy research. Although such measures have the advantages of simplicity and face validity, as pointed out by Hartmann et al., (1977), they have important drawbacks. Thus, they may be susceptible to retrospective falsifcation and selective recall, may be unduly influenced by the patient's level of adjustment either at the start of treatment or at its end, and reveal nothing about the basis for the clinician's judgement which would provide the information required for replication. Parents' reports, often relied on in evaluation of child treatment, may be open to the same criticisms (Schnelle, 1974). An important drawback of both therapist and parent reports is that they may be biased by positive expectations of treatment effects. In the present study, different vantage points and domains of functioning were tapped by seeking information from teachers, the children themselves, and their peers, through a variety of measures. It is evident that change data can provide a complex intermeshing of 'real' improvement, value orientations towards change of the different judges involved, and elements of measurement unreliability. It is the researcher's task to evaluate the relative contribution of each.

(iv) <u>The significance of change</u>

While the above measurement issues are undoubtedly important in the assessment of outcome, there are broader issues which need to/

to be considered in an adequate evaluation. Kazdin and Wilson (1978) provide a useful discussion of a number of these issues. In pointing to the inadequacy of statistical differences between groups as a basis for judgements about therapies they raise several additional questions. First, if change has occurrd, how important is it? Is it of clinical as opposed to statistical significance, in enhancing the client's everyday functioning? Although the inappropriateness of statistical significance has been widely discussed, positive objective alternatives for evaluating clinical change have not been readily available and satisfactory criteria may be hard to establish. Kazdin (1977) elsewhere suggests that in some cases even a dramatic change may not be clinically signif-icant. He gives the example of reducing self destructive behaviour in an autistic child from 80% to 40% of the times observed. Even this large reduction is of questionable clinical significance given that self destructive behaviour of any rate is maladaptive. Kazdin and Wilson (1978) suggest that the use of normative groups may help to establish clinical significance in a process of 'social validation'. Thus, a judgement may be made as to how an improved level of functioning in a treatment group compares with that of a group which is selected to be representative of the population.

A second factor to which Kazdin and Wilson direct attention is that of breadth of changes which may occur over and above alter-ation of the original problem for which help was sought. Treatment effects commonly extend beyond direct targets of treatment (Sloane et al., 1975) so measurement procedures need to cater for this. Of/

Of related concern are the 'side effects' of treatment which may lead to one treatment being preferred to another. For example, Ayllon et al., (1975) found that reinforcement techniques and Ritalin were equally successful in controlling hyperactive behaviour in classrooms, but the drug was found to suppress, whereas the reinforcement procedure accelerated, academic performance. Another 'side effect' may involve client satisfaction with therapy which could entail differential attrition rates across therapies.

(v) Assessing the maintenance of treatment

The need to ascertain the durability of treatment effects is widely accepted in therapy research. As Rutter (1982) points out, it is not enough to assess outcome at the end of treatment; it is essential also to determine whether the benefits remain some years later. Why is follow-up necessary? First, improvement that is apparent during or just after treatment may not persist once treatment has ended. Kent & O'Leary (1976), for example, found that children with conduct problems treated by behavioural means showed much greater improvement than a randomly assigned no-treatment control group. However, nine months later, this significant difference had evaporated. Follow-up is an especially important requirement for the behavioural field as a whole, given the unfavourable discrepancy often found between status at termination and status at follow-up (Mash & Terdal, 1977).

Second, follow-up is also important to evaluate benefits which may emerge only after treatment is completed. Sargent (1960) suggests that patients leaving therapy 'uncertain and discontented with its/.....

its results may realize benefit many months or even years after therapy has ended' (p. 495). Wright et al., (1976), as already noted, have documented such gains in their analysis of six controlled studies of individual child psychotherapy, with more gains evident at follow-up than at termination.

A third reason for undertaking follow-up evaluations is that, given the fluctuations which may occur after treatment is ended, it is possible that treatments which are effective in the long term are not the ones which are effective in the short term. Rutter (1982) cites a study by Craighead et al., (1981) to illustrate this possibility, in which behavioural treatments proved superior to medication for obesity in the long term, despite the greater potency of drugs in inducing initial weight loss. In a study comparing several different treatments, such as the present one, follow-up is therefore an important consideration.

Despite the general acceptance of the importance of follow-up data, the long term effects of many procedures remain a matter of speculation, because the necessary information has not been gathered. There is a paucity of follow-up data in child therapy research and as Mash (1975) has pointed out, when follow-up assessments are included, they are typically brief and unsystematic. In an analysis of the contents of four major behaviour therapy journals for the year 1973, Cochrane and Sobol (1976) found that only 35% of studies where follow-up was appropriate actually included follow-ups, and less than a third of these took place more/...

more than six months after the end of therapy. Keeley et al., (1976) in a similar analysis of three behavioural journals for the years 1972-73 estimated that only about 12% of studies provided follow-up data of more than six months' duration.

The collection of adequate long term follow-up data poses many problems which deter researchers. Of these, perhaps the most important is that of sample attrition. With longer follow-up intervals, factors such as subject mobility, family instability and name changes by female subjects make it increasingly difficult to contact former clients and their families for assessment (Hartmann et al., 1977). A number of reports illustrate this problem of attrition over time very strikingly, pointing to the difficulties of retaining a representative sample. (They also illustrate the difficulty in ensuring that a representative sample <u>enters</u> treatment.) Thus, Patterson (1974) indicates that of 35 referred families on which baseline data were collected, 27 completed at least four weeks of intervention, 21 a one-month follow-up, 18 a six-month follow-up and 16 a one-year follow-up. So approximately a quarter of the families dropped out prior to intervention, and little or no follow-up data were available for about half the sample. In a similar vein, of the 64 families seen by Johnson and Christensen (1975), for at least one intake interview, 22 remained for extended treatment, and 14 took part in follow-up assessments. Thus, follow-up data were based on only 21% of the initial sample. These authors also found that families remaining in the programme were rated as more co-operative by their families during treatment. As Gottman and Markman (1978) suggest, only 'satisfied customers' may/...

may return, increasing the possibility of bias in the evaluation. A further difficulty in gathering follow-up data is the lack of guidance or clear rationale as to when and how often follow-up should be conducted (Mash and Terdal, 1977). Decisions about follow-up intervals, according to these authors, require consideration of predictions for outcome based on available knowledge, and practical costs. Robins (1972) suggests that, on the one hand, the interval between treatment and follow-up can be too brief to allow an assessment of the stability of treatment effects, or, on the other hand, it may be so long that specific treatment effects are confounded with those of intervening life experiences. A problem with lengthy follow-up periods is that theorists who stress the reciprocal interaction between a person and his environment point to the complexity of analysing and controlling for influential factors in clients' lives long after treatment has ended.

With regard to frequency of follow-up, single probes are not adequate, given that the course or 'trajectory' of change may vary widely as a function of different treatments, and limited measurement probes may yield quite misleading impressions (Gottman and Markman, 1978). Multiple follow-up points can therefore provide valuable information on trends in change.

(vi) Conclusion

What conclusions may be drawn from this Chapter? Where the review of outcome is concerned, firstly, there is no satisfactory evidence that psychotherapy produces greater improvement than that which occurs in untreated control groups. This conclusion, however/

however, has been derived from the much criticised method of pooling studies, and it is also weakened by the absence of a commonly accepted estimate of the rate of change in the absence of treatment, which can serve as a reliable standard of comparison. A similarly inconclusive outcome is recorded for behavioural studies and those attempting treatment comparisons. Methodological limitations are again evident in both these categories of studies, and the scarcity of studies in the latter category reflects how little attention has been given to studying the comparative effectiveness of treatments with children. The firmest impression, perhaps, is that of the ubiquity of inconclusive findings, generated largely by methodological difficulties.

What, therefore, can be drawn by way of recommendations from the review of methodological issues in the second part of the Chapter? The following are the principal recommendations. First, untreated controls are necessary, in order to control for changes in the absence of treatment. These may occur as a function of 'spontaneous remission', developmental changes, or children's natural tendency to 'grow out' of problems. Second, cases should ideally be randomly assigned to treatment and control cases. Given the difficulties associated with defectors from treatment, these 'naturally occurring' control groups should be avoided. Third, where randomization does not ensure initial matching, differences need to be controlled for statistically. Fourth, where change is being assessed, global estimates should be avoided, and reliable measures need to be sought. Multiple measures of change/

change are preferable to single measures in order to reflect the different vantage points of observers, and the client himself, and different domains of functioning. Fifth, observed changes require to be assessed for clinical as well as statistical significance. Sixth, follow-up data should be obtained in order to assess the durability of changes and to examine possible shifts in comparative effectiveness. Several follow-up points are necessary if trends are to be identified.

This seems a daunting list of requirements, and it is, of course, a selective one. It can, however, function as a guiding set of recommendations, for the present study. As a survey of the therapy effectiveness literature quickly shows, there are many constraints, including time, manpower and resources, which conspire against the researcher's meeting any set of ideal methodological requirements. The present study was not immune from such difficulties.

CHAPTER 4

REVIEW OF THE LITERATURE: TREATMENT INTERVENTIONS IN EDUCATIONAL SETTINGS

REVIEW OF THE LITERATURE: TREATMENT INTERVENTIONS IN EDUCATIONAL SETTINGS

Introduction

The principal concern of the present chapter is to review the outcome of behaviour modification interventions within educational settings. The review will be restricted to studies involving maladjusted or emotionally disturbed children, and will not cover interventions with mentally handicapped, delinquent and psychotic children. Earlier chapters have provided a background against which these studies need to be viewed. Thus, issues in the nature and assessment of the maladjusted have been considered, the theoretical underpinnings of behaviour modification, as one intervention approach, have been examined, and attention has been given to the wider background of treatment outcome and research methodology in work with children.

This chapter will also present findings from the research literature which are relevant to the two treatment comparison conditions - parent counselling/teacher consultation and group counselling. Given the status of these interventions in the present study as comparison conditions, rather than as treatments to be fully explored in their own right, the review will not be as extensive as that for the main intervention, behaviour modification.

A. Behaviour Modification Studies

Introduction

The behavioural programme in the present study had as its principal objectives: (a) the modification of social behaviour deemed inappropriate by the teacher in the classroom, (b) an increase in task-related behaviour and (c) enhancement of the interpersonal functioning of those children who were isolated or rejected within the classroom group. These objectives reflect a focus on what might be seen as the correlates of 'maladjustment' in the classroom context, as suggested in Chapter One. In general, these objectives may be viewed as being designed to assist the behavioural, academic and social functioning of the children through classroom-based interventions. This section will examine the literature relating to those areas..../

areas. The study was also concerned to assess changes in other spheres, including ability, reading achievement, personality and school-related attitudes, and relevant behavioural interventions in these areas will also be reviewed. Following consideration of the impact of these interventions, and to the extent to which their effects are generalised, some of the issues involved in the training of teachers in behaviour modification approaches will be examined. In addition to the above concerns, two additional features of the study were of special relevance in determining the areas of work to be highlighted in the literature review. These features were, first, the involvement of children of secondary school age in ordinary school settings, and, second, the use of a group design, with untreated controls and treatment comparison groups. Studies in the secondary school setting, employing group designs, with control or comparison treatments, will therefore be given special emphasis. However, the ordinary secondary school setting has been relatively neglected in behavioural research, with most interventions having been conducted in primary schools or in special settings. In addition, the development of behavioural work in the classroom owes a great deal to studies employing other kinds of experimental design, such as reversal or sequential procedures with single subjects or within small groups. Findings from this wider context of behavioural research in educational settings will therefore also need to be considered in describing the principal issues in the area.

Over the period of time since the treatment approaches in the present study were formulated, developments have occurred in a number of areas. First, there has been increasing concern with the applicability of behavioural procedures within ordinary classrooms. This issue has been examined in the Introduction where its influence on the design of the study was discussed. Second, there has been a trend towards interventions which focus on aspects of academic functioning and performance as well as on social behaviour in the classroom. The issues underlying this trend will be examined later in this chapter. Third, increasing attention has been given to 'cognitive-behavioural' approaches, with a growing recognition within the field of behaviour modification of the role of attitudes, beliefs, expectancies and attributions.../

attributions in understanding behaviour and behaviour change. This approach subsumes developments in the areas of self-reinforcement, self-instruction and self-control all of which have been explored in an educational context. These areas of investigation will not be covered in this review since they represent a divergence from approaches stressing reinforcement by teachers, this being the emphasis of the present study. Possible developments and extensions of the present approach will be considered in the Discussion Chapter. Detailed reviews of cognitive-behavioural methods are available in Meichenbaum (1977), Kendall & Hollon (1979) and O'Leary & Dubey (1979).

A noteworthy feature of the present review is that the studies are predominantly American in origin. Although a number of British studies have been published since the early seventies, most of them, as pointed out by Merrett (1981) in a recent review, have been presented in the form of studies without experimental controls. They have also been mostly confined to primary school settings, so fall outside the main focus of the present review. Although behavioural interventions have been somewhat slow to develop in this country, at least relative to the American context, more and more reports are appearing and important initiatives are being taken in teacher training (Wheldall & Merrett 1982). Recent publications by Wheldall (1981) and Harrop (1983) provide an overview of British work in a variety of educational settings.

(i) Interventions for inappropriate behaviour

(a) Introduction

Some clarification is needed at the outset for the concept of 'inappropriat behaviour'. In the present study, the teacher was taken as the arbiter of what would be deemed 'appropriate' or 'inappropriate' behaviour in the classroom. Although this would undoubtedly vary across teachers and subjects, 'inappropriate' can be broadly taken to mean those behaviours which would be seen as interfering with teaching activities or impeding the children's learning.

This would include disruptive and aggressive behaviour, disobedience, inattention and so on. 'Appropriate' behaviour on the other hand, might include listening to instructions, contributing to discussions, sustained concentration and persistence with tasks and so on. Critics have suggested that the adoption of teachers' points of view and the focus on eliminating behaviour teachers consider disruptive is to support an educational system which may well be unsatisfactory (Winett & Winkler, 1972). Adjusting individuals to a poorly designed environment may be a disservice. However, it can also be argued that there are merits in the above approach in that these 'appropriate' forms of behaviour have important implications for children's academic progress and their effective functioning in a task-orientated group. Furthermore, behavioural approaches do attempt to change children's school environment, principally by modifying their teachers' management methods, making them, hopefully, more positive and less coercive in orientation. Other models of intervention may, of course, take a quite different stance, emphasising, for example, a more co-operative approach to rule-setting (e.g. Glasser, 1977) or explicitly seeking to understand 'disruptive' behaviour from the pupil's point of view (e.g. Marsh et al., 1978).

(b) Studies emphasising social reinforcement

Probably the most widely used procedure in classroom behaviour modification have involved variants of social reinforcement, with teachers systematically employing methods such as praise or approval in relation to desired behaviour, and disapproval or withholding of attention in relation to behaviour considered inappropriate. This emphasis on social reinforcement is not surprising, given that such methods exist to some degree in all teachers' management repertoires. In persuading teachers that this approach is a worthwhile one, it is indeed important to show that the use of social reinforcement largely involves a re-organisation or re-structuring of something that they already do, rather than the introduction of a novel and esoteric set of procedures.

However,..../

However, while all teachers use elements of social reinforcement, this does not mean that such procedures are always used appropriately or effectively. Praise may not have a sufficiently precise focus in the child's behaviour, for example, or it may occur in a non-contingent fashion, without having a clear relationship in time to the desired behaviour. Another consideration is that praise may be used appropriately in these respects, but be quite insufficient in quantity, or outweighed by negative methods such as criticism, disapproval or reprimands. It is usually recommended within behavioural programmes that the balance between positive and negative forms of teacher attention should be heavily weighted towards the former. Madsen et al., (1970), for example, suggest a positive to negative ratio of 4:1 as the optimum.

In relation to this last point, there is evidence that many classroom settings are far removed from this ideal. Reference has already been made in the Introduction to data suggesting that the rate of teacher approval diminishes after first and second grades and is thereafter consistently exceeded by the use of disapproval (White, 1975; Madsen & Madsen, 1973). This imbalance in favour of negative procedures has also been found recently with 11 to 13 year old groups of 'problem' children in guidance units set up to convey special help (Thomas et al., 1978). Although these studies have all been conducted in American classrooms, the British context is not dissimilar: Rutter et al., (1979) found that secondary teachers in London tend to reprimand twice as often as they praise.

A common explanation of these findings is that disapproval and reprimands, by often tending to terminate misbehaviour immediately, serve to reinforce the teacher in this behaviour, so their use tends to continue. However, this seems to be an unfortunate 'trap' since this positive effect tends to be temporary and in the long run, continued use of high disapproval rates may..../

may prove quite ineffective. Teachers' negative attention to undesirable behaviour may have quite the opposite effects to those intended (Madsen et al., 1968; Thomas et al., 1968). In the Madsen et al. study, in which a teacher was trying to control children getting up out of their seats, an increase in 'sit-down' commands actually led to an increase in out-of-seat behaviour. The giving of approval on the other hand, may not have had the same immediacy of impact for teachers, and some teachers feel that appropriate behaviour - 'what children should be doing anyway' - deserves little recognition (Thomas et al., 1978).

The fall-off in teacher approval rates after the early years may be related to the possibility that children make obvious progress in learning more slowly and find school learning increasingly less rewarding. Under such conditions the teacher may find her own efforts progressively less rewarding, and her interactions become increasingly less positive. A shift to more positive forms of control, with emphasis on increased use of approval, may therefore require a not insubstantial change in teacher behaviour. The procedures of social reinforcement may appear deceptively simple, but the extent of change required, on the teacher's part, should not be minimised - teachers have their own established patterns of behaviour, their own reinforcement history, which cannot be easily discarded and laid aside. Hall, Lund and Jackson (1968) for example, found that two of the five teachers in their study were unable to carry out systematic praise and ignore procedures. Ward (1971) in the U.K. found similar difficulties on the part of one of the three teachers with whom he was working. However, many studies are now available which point to the benefits, for both teachers and children, which may accrue from the skilled and sensitive use of social reinforcement. Before examining some of these findings, a number of issues in the early development of behavioural work in educational contexts need to be considered.

A major concern of the present study was the application of reinforcement in ordinary school settings. At the time of planning the study, however, while..../

while a number of examples of such interventions were available, many of the published reports had come from rather specialised settings which left questions about generalisability of findings unanswered. Thus, investigations were conducted in self-contained special classes for the handicapped (Birnbauer et al., 1965; Haring, 1968), adjustment classes for the emotionally disturbed (O'Leary & Becker, 1967), special laboratory or experimental classrooms (Hewett et al., 1969; Walker et al., 1971) or experimental nurseries (Allen et al., 1964; Hart et al., 1968). Many of the classes in these studies were small and the children relatively 'homogeneous' in terms of presenting similar kinds of learning and/or behaviour problems, so the same kind of programme could justifiably be extended to them all. Such settings, therefore, were rather different from the regular school situation where only one or two children in a class group of some 30 children might require help.

Much of this early work was also imbued with the 'unbridled optimism' that characterised the reports of applied behaviourists in the 1960s (Baumeister, 1969). This optimism was shown principally in the belief that failures to learn social or academic behaviours were failures of the therapists' ability to structure tasks and arrange contingencies appropriately. Researchers in educational settings were anxious to show that behaviour modification 'worked' in the same way that behaviour therapists had shown their procedures to work with clinical problems - for example, Patterson (1965) with a hyperactive child and Lovaas (1969) with an autistic child.

Rather ironically, as pointed out in the Introduction Chapter, important contributory factors in the success of such 'demonstration' models, may have been the special advantages that attached to the very context in which the studies were conducted, i.e. their research context. Kent & O'Leary (1976) have pointed out how such advantages may have contributed substantially to the positive outcomes reported. Thus, teachers engaged in such projects have received college course credits or stipends in return for their participation...../

participation (Becker et al., 1967; Kaufman & O'Leary, 1972; Walker & Buckley, 1972). The ready availability of research staff allows for immediate feedback to teachers when help is required and for continuous support and encouragement. Teacher motivation may also be increased by the possibility of joint authorships or registration for a higher degree. O'Leary & Kent suggest that advantages such as these, which are not routinely available in regular settings, may have contributed to an over-estimation of the value of child behaviour modification.

Despite such difficulties, many of which were apparent only with hindsight, these early pioneering studies provided a stimulus and model for subsequent investigators to follow and helped to lay the ground rules for the application of reinforcement procedures. Thus, studies in preschool settings illustrated the use of positive reinforcement (e.g. praise, attention, verbal approval) following desired or appropriate behaviour, in combination with either extinction (ignoring) of inappropriate or incompatible behaviour, or the delivery of a reprimand following the undesirable behaviour. These studies suggested that social reinforcement techniques could be employed effectively to increase walking and reduce regressive crawling (Harris et al., 1964), reduce crying (Hart et al., 1964) and develop more active motor skills (Johnston et al., 1966).

Similar techniques were then extended to older children in primary and secondary school settings. With this extension, we find applications in regular classroom settings. Madsen et al, (1968) for example, working in a public elementary school setting, chose two children from two separate classrooms, one with 29 children, the other with 20. These children were selected because they showed a high frequency of problem behaviour, such as non-compliance, inattention, and aggression towards other children. Following a baseline phase, three experimental components - rules, ignoring disruptive behaviour and praise for appropriate behaviour - which had been found..../

found effective in combination in an earlier study (Becker et al., 1967), were introduced sequentially. It was found that rules alone were ineffective in modifying behaviour, and the rules and ignoring phase was associated with increased disruption. When praise for appropriate behaviour was added to the two other elements, inappropriate behaviour fell from about a 70% baseline level to about 30%. Reversal to baseline conditions produced a rise in inappropriate behaviour to its original level, and when a final reinstatement of experimental conditions restored satisfactory levels, their functional role was demonstrated. While the authors point to the effectiveness of combined praising and ignoring procedures and suggest that praise for appropriate behaviour is probably the key to effective classroom management, it must be pointed out that their experimental design, which cumulatively combined procedures, rather than isolating them, makes clear cut conclusions difficult. Methodological niceties aside, this proved to be a seminal study, possibly not only because substantial changes were observed, but also because a fund of detailed information was supplied concerning the practicalities of defining and observing behaviour, and applying the experimental procedures. Numerous other studies have successfully exploited the systematic use of teacher attention to increase study, on-task and desirable social behaviours (e.g. Hall, Lund et al., 1968; Hall, Panyan et al., 1968; Wasik et al., 1969; Broden et al., 1970; Breyer et al., 1971; Petersen et al., 1971). Other studies have focussed on the reduction of disruptive behaviours, such as noisy behaviour in class, talking-out, tantrums and aggression (Thomas et al., 1968; Ward & Baker, 1968; Hall,Fox et al., 1971; Lates et al., 1971). These studies all emphasised the combined use of praising and ignoring procedures.

However these procedures may not always be adequate. 'Ignoring' may not be sufficient to extinguish inappropriate behaviour and stronger measures, such as disapproval, may be necessary. Although the introductory comments in this section pointed to possibly undesirable effects of teacher disapproval, it.../

it should be noted that some studies have demonstrated that it can be effective when used in conjunction with praise and approval. McAllister et al. (1969), for example, working in a secondary school setting achieved significant improvement in controlling children's talking-out and turning-around with a combination of praise to the class as a whole for appropriate behaviour and disapproval to individuals every time they talked out or turned around. Thus, selective use of disapproval may be beneficial, and such negative forms of intervention clearly cannot be ruled out when the learning situation is being disrupted, or there is risk of a child being hurt as a result of aggressive behaviour. In this connection, Madsen et al. (1970) have suggested that it is the ratio between positive and negative methods that is important. As already noted, they recommend a 4:1 positive to negative ratio as the optimum, but no data are presented to support this.

Under ordinary circumstances, 'negative' teacher attention usually takes the form of a loud reprimand which will be heard by all the children in the class, as well as by the child at whom it is aimed. Not only does this distract other children, and possibly disrupt their activities, but it also singles out the miscreant for special public attention, giving him the spot in the limelight which tends to perpetuate rather than eliminate misbehaviour. Constant loud nagging may also create feelings of tension and resentment which militate against settled work. O'Leary et al. (1970) investigated whether teachers might be able to use an effective alternative to loud reprimands which did not have undesirable characteristics. Their alternative was a soft reprimand, a verbal corrective statement, delivered in such a manner that only the offending child was able to hear it. They contrasted loud and soft reprimands with two children in each of five classrooms. The subjects were aged 7-8 years and were selected because of high rates of disruptive behaviour. An ABAB reversal design was employed, with the A phases consisting of teachers using their customary loud reprimands, and B phases consisting of reprimands delivered at the same rate but in a soft manner. Each condition was in operation for one month. Soft reprimands proved.../

proved more effective than loud reprimands in reducing the mean frequencies of disruptive behaviour. However, the intervention did not work equally well with all children (two children in one class increased their disruptive behaviour with soft reprimands, leading to a cessation of these procedures) nor did all the teachers implement the method satisfactorily. One teacher viewed soft reprimands as a sign of weakness and used them inconsistently. She reported finding it too strenuous to walk over to the child's desk to whisper a reprimand and reverted to her 'natural tendency to shout like a general!'

While these findings suggest an important role for soft reprimands, it is stressed by the authors that they should not be viewed as an alternative to praise but as a procedure to be used in conjunction with it. Such a combination was in fact incorporated in a structured treatment programme prepared in the form of a manual by O'Leary and his colleagues (Kent et al., 1976). For present purposes, it is to be noted that soft reprimand procedures have not been explored in secondary school settings. One of the reasons advanced for their effectiveness is that through being less noticeable to other children, peer attention and reinforcement are reduced. With older children, who are possibly more sensitive to peer attention, the procedures may have quite different implications. Findings with younger children may not therefore be transferrable in a straightforward way. Further research is required to clarify more general questions as to why and for whom the procedures are effective.

In this context it is instructive to recall that what is 'reinforcing' for one child may not be so for another. Some may thrive on disapproval, and may indeed resent praise given in public if, as Marland (1975) suggests, they see authoritative commendation as a threat to their reputation. With such children, praise may well be effective, but only when given in private. That teachers have repertoires characterised by high rates of reprimands was suggested by Workman et al., (1980) as a reason for recommending the use..../

use of reprimands, at least initially, in conjunction with praise for appropriate behaviour. They found that more rapid and slightly greater improvement was achieved with a praise-soft reprimand combination than under a praise-ignore combination. In contrast, however, improvements in behaviour diminished less rapidly in reversal phases following the praise-ignore condition. This latter finding could have important implications for the maintenance of change obtained with these procedures, but it requires replication.

As O'Leary and O'Leary (1976) have pointed out, the effectiveness of experimentally manipulated teacher attention has been demonstrated with a wide range of subject populations in a variety of contexts, ranging in age from preschool (Schutte & Hopkins, 1970) to groups of 17 year-olds in secondary classes (McAllister et al., 1969), from classes of three children (Zimmerman & Zimmerman, 1962) to classes of 39 (Hall, Lund & Jackson, 1968) and from normal classes (Madsen, Becker & Thomas, 1968) to classes for the retarded (Hall, Fox et al., 1971). Some studies may focus on only one child in a class (e.g. Kirby & Shields, 1972), others may include the behaviour of an entire group (Hall, Panyan, Rabon & Broden, 1968). Where the behaviour of an entire class of pupils is the focus of intervention, it is clear that the teacher cannot attend to individualised treatment prescriptions for each and every child, as this would impose intolerable strain. The usual approach in these situations is to deal with common problem behaviours, such as talking when pupils should be working quietly, or deliberate noisy behaviour, and the treatment procedures are applied right across the classroom group, to any pupil displaying these behaviours. In the present study, something of a compromise was sought, with social reinforcement procedures being applicable to the entire class on a general basis, but with a special focus on the selected 'target' children within the class. Where point and token systems are concerned, a number of group contingency approaches have been developed to influence the behaviour of entire class groups. These are reviewed by Litow & Pumroy (1975).

Despite.../

Despite the wide range in applications of social reinforcement procedures, it is evident that the bulk of interventions have taken place in elementary or primary school settings. This may not be entirely unrelated to White's (1975) finding of a decrease as children grow older in teachers' use of social reinforcement, i.e. teacher behaviour with older children may be perceived by researchers as less amenable to change, or, as a 'baseline', it may be seen as too far removed from the ideal. Alternatively, children may be oriented more to attention from peers than from teachers. Whatever the reasons, there is a relative dearth of studies with secondary school age children. However, there is evidence that the secondary school age child can also be influenced by social reinforcement.

In a rare example of work conducted in a British rather than an American setting, Wheldall & Austin (1980) report a study conducted with twenty-five 14-15 year old children in one secondary class in the West Midlands. They employed a somewhat unusual selection procedure, asking the head teacher to consult with his staff in order to identify a 'particularly disruptive' class. The basic criterion for selection of the class was that most teachers, when faced with it, would say or feel 'not that lot again' or, when finished with it, would say or feel 'thank goodness that's over for another week'. The authors ended up working through the head teacher with this class, because none of the other teachers would volunteer. With the head teacher teaching mathematics, baseline observations established that the average percentage on-task behaviour was around 55%. It was noted that most off-task behaviour occurred when the quicker members of the class had finished the task set them, so the first intervention involved asking the teacher to set extra problems. This simple procedure resulted in a rise of average on-task behaviour to.../

to 69%. Next, they introduced a condition which alternated between (a) rules, praise and ignoring procedures and (b) a 'timer-game' with points awarded for on-task behaviour when a chime sounded on a variable interval basis. Points gained earned free time during a maths lesson. On the days when 'rules, praise and ignoring' strategies were in operation, on-task behaviour rose to over 80%, with a further rise to over 90% when the 'timer-game' was in operation. The authors comment on the gratifying effect of praise procedures with this population, although, strictly, the sequential design of the study precludes an evaluation of the impact of praise in isolation from the other procedures.

Cormier and Wahler (1973) undertook an examination of the effects of teacher contingent and non-contingent (random) social reinforcement (praise and/or attention) on the classroom behaviour of economically disadvantaged adolescents. The study was conducted with six classes of 14-year-olds in a junior-senior high school. Six teachers volunteered to participate in the study, and each identified three pupils in his class as being either disruptive or not motivated to do assigned work. 'Non-target' children were randomly identified and observed in each class as a means of assessing any carry-over of effects. The length of time necessary to demonstrate changes in behaviour was also explored. Instructions were given to the teachers on ignoring inappropriate behaviour, contingent praise and attention and non-contingent praise. There were six experimental conditions, with long and short time periods of contingent, non-contingent and control conditions. During control conditions teachers were instructed to reinstate the baseline conditions. The long time periods were eight days and the short four days. The six conditions were sequenced randomly. It was found that both contingent and non-contingent reinforcement were associated with significant increases of relevant and significant decreases in inappropriate behaviours when compared to control conditions. An intriguing finding is that no significant differences between contingent and non-contingent conditions were found. Non-target subjects showed favourable changes also, under both contingent and non-contingent.../

non-contingent conditions. Long time periods produced the more favourable results, for target subjects only.

This report is available only in a brief thesis abstract, and descriptive details are insufficient to allow an adequate evaluation. Alongside the encouraging findings on the responsiveness of adolescents to social reinforcement, the impact of non-contingent reinforcement is unexpected. Most writers in the field stress the importance of making teacher attention contingent upon desired behaviour, yet Cormier's findings suggest that this close relationship may not be necessary. This finding is unsupported possibly because contingent versus non-contingent comparisons have not been carried out. Typically, investigators compare contingent reinforcement with 'baseline' conditions which usually involve disapproval and reprimands rather than high rates of random approval and positive attention. The question of whether simply increasing teachers' positive attention is in itself worthwhile therefore deserves further examination. Sajwaj and Dillon (1977) have suggested that attention may serve as stimulation which could have direct or indirect behavioural effects independent of any contingent usage. They indicate that work with infants who are failing to thrive points to dramatic improvements in physical and behavioural development coincident with an increase in non-contingent attention. Similar stimulation effects may well operate with older children in different settings.

Two further studies, by McAllister et al., (1969) and Marlowe et al., (1978), also showed positive effects with social reinforcement with older pupils. Their work will be described below in the section dealing with studies employing control/comparison group designs.

What is it that makes teacher attention so effective? Apart from its reinforcement or incentive value for children, it may also have an informational function, in cueing the child to behave in a particular way, especially when it is specific ("I'm pleased that you are asking more questions in class, Billy") rather than vague ("You've been a good boy today").

This purely informational function is suggested by Drabman and Lahey (1974) in reducing the disruptive behaviour of a ten-year-old girl, simply by giving her evaluative feedback in the form of a rating every 10-15 minutes. Significant reduction in levels of disruptive behaviour occurred with provision of feedback in an ABAB design.

The opportunity for observational learning on the part of other children also exists, so that the child who is praised acts as a 'model' of desired behaviour for others in the classroom. Drabman and Lahey, in the above study, found that the target subject's classmates also reduced their disruptive behaviour, although the contingencies did not apply to them. Broden et al., (1970) obtained a similar effect with two disruptive six-year-olds seated at adjacent desks. When the teacher increased her praise for one of the boy's attending behaviour, the second boy's attending also improved. Kazdin (1977a) demonstrated similar effects with retarded children and suggested that approval to one individual serves as a discriminative stimulus to another individual that his or her behaviour is likely to be reinforced. Delivery of approval therefore provides cues to non-reinforced individuals to behave well. Kazdin suggests that when the anticipated reinforcement does not materialise, the non-target subjects' appropriate behaviour should decline. However, this effect is not observed in either Kazdin's own study or in others over several days, and findings for the longer term are not available.

Whatever is likely to prove the most appropriate interpretation of these effects on non-target subjects, it is of interest to speculate that these findings might also constitute evidence for suggestions that non-contingent reinforcement may also have an instrumental role in behaviour change, as in Cormier & Wahler's (1973) experiment, simply by increasing levels of teacher approval. In the experiments showing changes in non-target..../

non-target children, these children would possibly have been aware of increased levels of teacher approval but, since this was not directed specifically at them, it would be experienced as non-contingent.

(c) Studies emphasising token reinforcement

One fundamental drawback in the use of social reinforcement is that it may not be sufficiently powerful with some children to alter their behaviour. Teacher praise may, as already noted, indeed be aversive in some cases. In such situations the use of token procedures, with a range of attractive back-up reinforcers (sweets, access to preferred activities, choice of toys, etc.) in addition to systematic use of teacher attention, is likely to be more effective. It is worth stressing that it is desirable to make use of tokens in addition to and not instead of teacher praise and attention. Tokens and the rewards to which they lead are essentially props which facilitate behaviour change. The intention of such a programme, and the task facing the person administering it, is to bring the resulting behaviour change ultimately under the control of more naturally-occurring influences in the child's environment, and appropriate use of teacher attention has a critical role here.

Token programmes usually involve three factors: (1) a set of instructions to the class about behaviour that will be reinforced, (2) a means of making a potentially reinforcing stimulus, usually tokens (e.g. check marks, chips, points) contingent upon behaviour and (3) a set of rules governing the exchange of the tokens for 'back-up' reinforcers, such as privileges, material rewards or access to preferred activities (O'Leary & O'Leary, 1972).

Kazdin & Bootzin (1972) list the following advantages of token programmes: (1) consequences for a response can be provided immediately, (2) they bridge the delay between target responses and back-up reinforcers, (3) they can maintain performance over extended periods of time when the back-up reinforcers cannot be administered, (4) they allow sequences of responses to be reinforced without interruption, (5) the variation available in back-up reinforcers reduces..../

reduces satiation effects and (6) they permit the use of the same immediate reinforcers for individuals with preferences for different back-up reinforcers. That token programmes are potentially more powerful than teacher attention alone was demonstrated in a study by O'Leary et al., (1969). Working with seven problem 8-year-olds who were prone to wander around, hitting other children and making noises, they found out that a combination of rules, altered classroom structure, and 'praise and ignore' techniques was generally ineffective in reducing disruptive behaviour. Significant improvements were recorded only with the introduction of a token programme. Main & Munro (1977) conducted a modified replication of this study by applying a series of techniques including individualised instructions (structure), praising and ignoring, tokens, and contingency contracting with junior high school students. They demonstrated that all four procedures in combination produced the highest on-task levels, and that these levels were maintained one month later when only contracts and praise/ignore procedures were in effect.

It is to be noted, however, that while token procedures may have more impact than social reinforcement alone, they do not necessarily have their effect in isolation from teacher attention. It has indeed been suggested that the use of tokens may be important precisely because it provides a way to structure and prompt teachers' attention to children rather than as a unique incentive in itself (Liberman et al., 1975). Thus, it has been noted that participation in a token economy was associatedwith a reduction in reprimands and disapproval, a reduction in attention to inappropriate behaviour and an increase in attention and approval to appropriate behaviour (Breyer & Allen, 1975; Chadwick & Day, 1971). Breyer & Allen in fact found that they were quite unable to increase a teacher's praise rates and to decrease reprimands by instruction and feedback, and secured the desired changes only with the introduction of token procedures.

Many reports point to the impact of token procedures in increasing task attention (e.g. Bushell et al., 1968; Broden et al., 1970; Sulzer et al., 1971; Ferritor et al., 1972) as well as reducing negative, disruptive and deviant behaviour (e.g. O'Leary & Becker, 1967; Carlson et al., 1968; Kuypers et al., 1968; Walker & Buckley, 1968; Wolf et al., 1970). A comprehensive review has been provided by O'Leary (1978). Despite the apparent advantages of token procedures, their use in the classroom may seem alien to many, and questions have frequently been raised as to their relevance to the regular classroom setting. Concern has perhaps been greater where tangible items, such as sweets, toys or comics, rather than naturally-occurring activities, are scheduled as rewards. What often appears to be objected to, in essence, is the systematic and planned application of rewards in the classroom. Tokens represent a very explicit and concrete form of reward and token programmes may accordingly more easily attract the range of criticisms that might be directed at behaviour modification approaches. A frequent objection is that such approaches constitute bribery. Although the customary immoral or corrupt connotations of bribery can hardly be attached to the objectives of classroom token programmes, the unease felt by many teachers and parents is not easily dissipated. It may be helpful in trying to meet these criticisms to point out that 'tokens' of a kind may already be commonly available in schools in the form of stars, house points, certificates, etc. A formal token programme may therefore differ from current practices only in its systematization of rewards. Despite the potential ethical and practical difficulties surrounding their use, McLaughlin (1975) suggests in his review of the literature that token systems can be adapted successfully to regular classrooms. He points to the importance of such aspects as ease of implementation and management, cost of running, compatibility with school and community attitudes, and pupil satisfaction in determining the viability of programmes.

Some objections have focussed on the potentially harmful effects of token procedures both on children who are excluded from programmes and also those..../

those who are involved in them. The former concern is that where token reinforcement is applied selectively to a few children in a class, the others may feel deprived, or perhaps misbehave in order to obtain reinforcers themselves. What evidence is available on this issue suggests that, on the contrary, the effects on other children may be favourable: Christy, (1975) showed that the use of tangible reinforcers with target children in a preschool setting resulted in behavioural improvement in unrewarded peer observers. Whether similar effects could be demonstrated with older and more socially sophisticated children is unclear. It does seem, however, that where reinforcement takes the form of teacher attention and approval, the effects on unrewarded peers is generally favourable (Broden et al., 1970; Drabman & Lahey, 1974; Kazdin 1977a).

The concern that token programmes may have deleterious effects on those involved in them has grown out of studies dealing with the effects of extrinsic consequences on intrinsic interest or motivation. Work by Deci (1971) with college students and by Lepper et al. (1973) with preschoolers showed that providing extrinsic rewards for performing an intrinsically interesting task led to a subsequent reduction in interest in the task and interfered with performance after the rewards were withdrawn. In what has become known as the 'overjustification hypothesis', one of the implications of attribution theory has been invoked, stating that if a person engages in a high-interest activity in order to obtain some particular reward, then this person is likely to infer that the behaviour is controlled by the reward and his interest in the activity will thereby decrease when the reward is no longer available. Critics such as Levine & Fasnacht (1974) have drawn on the above findings and their implications to make sweeping attacks on token reinforcement programmes and to point to their likely harmful effects. In their view, tokens may lead to short-term improvements but long-term decrements in behaviour.

Notwithstanding the flurry of publications in 'the great overjustification debate'..../

debate' (Franks & Wilson, 1979) there are a number of basic weaknesses in the criticisms of token programmes which have flowed from it. First, and most seriously, the argument based on the overjustification hypothesis is predicated upon external reward of high-frequency behaviours that are intrinsically reinforcing. Yet these are circumstances in which token reinforcement programmes would not normally be used. If a child was already intrinsically motivated and engaged in tasks at a high frequency, intervention would not be necessary. Second, findings within the overjustification paradigm may be of limited generalisability. They stem from investigations of performance on laboratory tasks which differ markedly from the usual behaviour reinforced in token economies. As Kazdin (1977b) has pointed out, these differences may be important in that many of the target behaviours in token economies have some influence or operate on the natural environment after they are performed. For example, mastering of an academic assignment after token reinforcement is likely to bear some rewards of its own. Thus, behaviour might be maintained after extrinsic consequences have been withdrawn. Another feature of these studies, which limits their generalisability, is that they have predominantly focussed on the effects of brief periods or of relatively few trials of reinforcement unlike the extended periods of reinforcement characteristic of token programmes. Where the latter conditions are more closely approximated, evidence for overjustification effects tend not to be found (Feingold & Mahoney, 1975.).

Although many of the criticisms based on the overjustification hypothesis may have been largely unjustified or misdirected, they have had some important effects. First, they have directed attention to the role of subjects' attributions, and the understanding that they develop of their involvement in token programmes. In this connection, it is interesting to note that Dollinger & Thelen (1978), in a study within the 'overjustification' framework, found that the more controlling tangible rewards they employed had a more detrimental effect on intrinsic motivation than verbal, informative or symbolic rewards. Second, they have led to greater acknowledgement that tokens should be employed as a last resort, where..../

where simpler and more natural strategies - such as social reinforcement - have failed. Third, they stress the need for attention to be paid to the issue of maintenance in behavioural research, and what happens when programmes are terminated.

In contrast to the paucity of investigations of social reinforcement procedures with children of secondary school age, there are many examples of applications of token reinforcement with these older children. Meichenbaum et al. (1968) worked in an institutional setting with ten female adolescents and, using a reversal design, achieved an increase in appropriate class behaviour when money was made contingent upon such behaviour. Kaufman & O'Leary (1972) and Santogrossi et al (1973) reported similar success with emotionally disturbed adolescents in psychiatric hospital school settings. In both studies, the introduction of token reinforcement was instrumental in reducing disruptive behaviour, and in the former study, it was also shown that maintenance of improved behaviour was possible for a short period when the adolescents themselves, rather than the teacher, evaluated their behaviour. Broden et al. (1970) with a special education classroom group of thirteen 13-14 year olds, increased on-task behaviour from the baseline level of 29% to 57% by the use of contingent teacher attention. When the token system, which involved receiving points which were exchangeable later for extra time at the lunch break, was employed, study behaviour increased to 74%. Blanchard & Johnson (1973) extended the use of token procedures into the ordinary school setting, working with five 'behaviour problem' 13-year-olds in each of two classrooms. Through the twelve weeks of this experiment, baseline conditions, in which the two teachers conducted their experimental classes in their own typical manner, were alternated on a weekly basis with six separate experimental conditions: ignoring inappropriate behaviour; approval for appropriate behaviour, disapproval of inappropriate behaviour, combined praise-and-ignore procedures; tangible individual rewards and punishments. Tangible rewards and punishments proved the most effective procedure.../

procedure in increasing appropriate classroom behaviour and decreasing disruptive behaviour, regardless of the teacher applying them.

These studies demonstrate something of the impact of token reinforcement approaches with secondary school age children, albeit with special populations and somewhat small numbers. Of greater relevance to the present investigation is the availability of several evaluations conducted in ordinary school settings which have included control groups in the experimental design. These will be discussed below in the separate section dealing with such studies.

(d) Studies emphasising punishment procedures

The procedures described in this section so far have been predominantly positive in orientation, and this is a fair reflection of the spread of methods in the field. Procedures within the category of 'punishment' have been investigated relatively little. O'Leary & O'Leary (1977) suggest that this has been because psychologists have espoused the 'legend' that punishment is an extremely ineffective means of controlling behaviour, and also for fear of being labelled as punishment advocates. It is often suggested that punishment of any kind may damage the relationship between child and teacher, especially if the punishment is seen as unfair and designed to humiliate (e.g. Fontana, 1981). In addition, the child may adopt strategies, such as untruthfulness, to avoid punishment. Where corporal punishment, especially, is concerned, an inappropriate model is presented to children of the acceptability of imposing penalties on those weaker than themselves. Despite these drawbacks it is suggested that punishment can be effective under certain conditions, e.g. when administered at the onset of a misdemeanour (Aronfreed, 1976), when it is consistently applied (Parke & Deur, 1972), and when there are alternatives to the punished response for which praise and approval are available (Clarizio, 1976). It is also to be noted that the forms of punishment which have been explored in classroom or child development experiments have been..../

been rather unlike the usually highly aversive forms implicit in everyday conceptions of the term.

Typically, classroom punishment procedures take three main forms (O'Leary & O'Leary, 1977). First, the use of disapproval and reprimands, which has already been discussed in the context of social reinforcement. Second, there is the procedure of 'response cost' or points loss within a token reinforcement programme. Examples of this latter procedure are studies by Meichenbaum et al. (1968) in which delinquent junior high school students lost points, exchangeable for money, for inappropriate classroom behaviour, and by Blanchard & Johnson (1973) where the loss of a variety of privileges was employed as a cost. Iwata & Bailey (1974) working with fifteen 10-year-olds found response cost and token procedures to be equally effective in improving classroom behaviour and output in arithmetic performance. The third method is 'time-out'. This procedure involves placing the child in social isolation contingent upon misbehaviour, for a short spell of time. It is commonly referred to as 'time out from reinforcement' but O'Leary & O'Leary (1976) point out that since this label makes an assumption about the nature of the environment that is rarely assessed, defining the process as social isolation is preferable. These authors also suggest that the procedure has been employed and evaluated in special rather than regular classrooms, and recommend that, in view of its potential unpopularity with teachers, it should be used very discriminately and usually only after other methods have been tried. In regular classroom settings, rigorous implementation of time-out procedures is often precluded by the unavailability of an adequate room which can be used for this purpose. In addition, the alternatives of standing in a corner, outside in the corridor, or going to the headmaster's or secretary's office often allow opportunities for inappropriate reinforcement. One successful application of time out in a secondary school setting is illustrated by Broden et al. (1970) who sent a student who was not responding to a points system to an isolated part of the principal's office in an attempt to..../

to reduce cursing and fighting. For a discussion of theoretical, legal and ethical issues, the reader is referred to Macmillan et al. (1973) and Gast & Nelson (1977).

(e) Studies employing control or comparison group designs

Most classroom behaviour modification studies have employed intrasubject designs, or within-group designs with small numbers, to demonstrate the impact of behavioural procedures or to compare the effectiveness of different kinds of procedures within the behavioural framework. Larger scale applications, employing control groups, have been rare, as have been attempts to compare behaviour modification with other treatments from an alternative theoretical framework. This section will examine the few available examples of such studies, with children of secondary school age. They include both social reinforcement and token reinforcement studies, and they have been drawn together here because of their relevance to the design of the present study. They are also located in regular secondary school settings.

There are two studies - McAllister et al (1969) and Marlowe et al. (1978) - which carried out an analysis of social reinforcement procedures. Marlowe et al., also included token reinforcement in a separate phase.

McAllister et al., (1969), in a study already mentioned briefly, provided one of the earliest examples of a behavioural approach at the secondary school level. An additional point of interest was the application of procedures to an entire class rather than one or two pupils. The objectives of the study were to reduce inappropriate talking and turning around while pupils were supposed to be engaged in academic tasks. The average age of the pupils was 17 years with 25 in the experimental class and 26 in the control class. The same teacher taught both classes, and with the control class, continued to teach in her customary manner. After a 27-day baseline, the teacher began to apply disapproval for inappropriate talking in the experimental class, with a direct, verbal, sternly given reproof. She used the pupils' names..../

names when correcting them. Turning around was not mentioned during this phase. Additionally, the teacher praised the whole class for periods of quiet. After 26 days of this condition, disapproval for turning around was introduced, with disapproval for talking being continued. This was continued for a further 8 days. A decrease in both these behaviours was attained. For talking, the baseline mean of 25% reduced to around 5%, while turning around reduced from 15% to 5% when intervention effects had stabilised. The control group means were comparable to baseline levels for the experimental group throughout the study.

The multiple-baseline design of this study presents compelling evidence for the effectiveness of social reinforcement procedures, with behaviours predictably changing during the specific intervention phase. It is perhaps surprising that talking and turning around did not display an interdependence which would have created difficulties for this particular experimental design. One usually thinks of talking and turning around as going together in the classroom so that change in one has implications for the other. Follow-up data were unfortunately not collected. The study nevertheless represents an interesting application of social reinforcement procedure with adolescents - albeit not with problems of serious proportions. Despite this, it did not establish a trend in work with this age group, as similar studies with primary-age children had done.

A more recent exploration of social reinforcement (amongst other procedures) is provided by Marlowe et al. (1978). This study is of additional relevance to the present project because of its comparison of client-centred counselling with behavioural approaches. The subjects in the study were twelve boys in a class of 30. Their ages ranged from 12 to 16 years because they had repeated one or more grades. They were selected because of their high rate of inappropriate classroom behaviour, as reported by the teacher and school principal. Following baseline observations of inappropriate off-task behaviour, the boys were divided into three groups, matched for average inappropriate behaviour...../

behaviour. These groups were then randomly assigned to one of three treatment conditions with 4 boys in each. One group received behavioural counselling consisting of 30-minute group sessions during which the counsellor emphasised the importance of on-task behaviour. The pupils received approval for positive behaviour while inappropriate behaviour was ignored, and they were also shown how to correct themselves when they had behaved inappropriately and how to credit themselves for appropriate behaviour. The second group of boys participated in 30-minute group counselling sessions based on client-centred principles. The emphasis here was on developing a warm and friendly relationship with the students, maintaining respect for them and their ability to solve their own problems, and showing sensitivity to their feelings. For both behavioural and client-centred regimes, the same individual served as counsellor.

The two counselling groups, and the third group, which as a no-contact control group received no counselling, experienced six experimental conditions. These were: (1) ten days of baseline, (2) eighteen days in which eight counselling sessions were conducted for the two groups, (3) five days in which appropriate behaviour was praised and negative behaviour ignored, (4) seven days in which token reinforcement was added to teacher attention with the entire class having the opportunity to earn tokens redeemable for doughnuts and soft drinks on the last day of the week, (5) four days in which baseline conditions were reinstated and (6) eight days in which teacher attention and token reinforcement were reinstated. Each counselling group participated in a total of 15 sessions and these were spread across phases (2) to (4). The 'no-contact' control group, it should be noted, was exposed to all the classroom-based procedures but not to counselling. The design allowed a comparison of the effect of counselling with no counselling, of teacher attention with and without counselling, and of teacher attention plus token reinforcement with or without counselling.

The results showed that teacher attention was highly effective in reducing pupils'..../

pupils' off-task behaviour, with a reduction to 40% of baseline levels across the three groups. However, the addition of token reinforcement proved even more effective with a reduction overall to 23% of baseline levels. Apparently the use of tokens served not only the expected reinforcement function but it also served as a reminder to the teacher to give verbal approval since this reached its highest frequency during the token phase of the study. Students who had received behavioural counselling reduced their off-task behaviour more quickly than students in either of the other two groups, and they showed significantly lower off-task behaviour than those in the other groups in all treatment conditions. It is suggested by the authors that the behavioural counselling had helped prepare the students for the teacher attention condition that was later implemented. They do not emphasise, however, that in the final phase of the experiment, <u>all</u> the students, including those who had received no counselling at all, reduced their off-task behaviour considerably below the baseline frequency. The no-counselling group reduced off-task behaviour from a 55% baseline level to 7.5%. Although the behavioural counselling group showed an off-task level that was significantly lower (2.9%) one may question the necessity of counselling at all.

There are several features of the design of this study which detract from the persuasiveness of the results. First, by employing the same person as counsellor in the two counselling regimes, as a means of controlling for counsellor personality and background variables, therapist skill and motivation in the two conditions may not have been balanced. Second, having subjects from all three conditions together in the same classroom may have created contamination. This may have militated against a treatment effect rather than facilitating it, but it does hinder a clear comparison of the three conditions. Third, the dependent measure of task-attention may have been biased in favour of the behavioural counselling group and against the client-centred approach. The overlap between the behavioural counselling experience and the subsequent classroom contingencies seemed considerable. The loose structure of the client-centred approach may well have been counter-

productive...../

productive (Ross, 1981). The comparison may not therefore have been a fair test of client-centred counselling. It should nevertheless be noted that in the final phase of the experiment, pupils in this condition showed a reduction in off-task behaviour from a 59% baseline level to 10%. This reduction would seem highly significant in practical terms, even although the behavioural group showed a lower level in the final phase. Finally, it is regrettable that the number of subjects was so small, and that no follow-up data were collected.

While the test of the two counselling methods is unconvincing, the impact of teacher attention, especially with the addition of token reinforcement, seems clear. Whenever teacher approval of positive pupil behaviour was low, off-task behaviour was high. Conversely, when teacher approval was high, inappropriate classroom behaviour was low.

Heaton et al. (1976) compared the effects of token reinforcement in one school (N = 14) with the outcome of traditional schooling for controls (N = 32) in another two separate schools. The sample of 14 year old children was selected as having presented serious discipline problems prior to entering the eighth grade. All the children had had two or more suspensions from junior high school for reasons other than smoking. In the token group, points were given contingently for starting, maintaining and completing assigned work as well as for social behaviour appropriate to the classroom. Points could be used to gain access to preferred activities such as pool, table tennis or cards, or to purchase items such as soft drinks and sweets, or to obtain early dismissal from school. The children were given a 'disturbing and disruptive behaviour slip' if they engaged in serious misconduct in the school. If two slips were received during one class period, the students were dismissed from the programme, sent home, and required to participate in a parent conference before re-entry into school. In addition to the above components, meetings were held periodically with the parents in order to provide feedback on children's performance at school, and to develop home..../

based reinforcement systems. Children in the control group remained in their regular school classes. These children were 'identified' to school personnel as being highly vulnerable, with a high potential for school drop out, low achievement and continued deviant behaviour. The programme was in operation for the full school year.

The treatment group did better than controls in a number of areas. First, in terms of 'holding power' the treatment group had a significantly better attendance rate (74% versus 56%) and fewer withdrawals from school. Fourteen control subjects were withdrawn from school for difficult or disruptive behaviour, compared with one treatment subject. Second, the treatment group had less in the way of misbehaviour problems, with fewer referrals to the school principal for misconduct, and fewer suspensions from school. Third, the treatment group demonstrated a greater gain than controls on a reading measure but differences in arithmetic and spelling were not significant.

Amongst the difficulties in evaluating this study, the authors point to a confounding of curriculum procedures and behavioural procedures. Although a special curriculum distinct from the control curriculum was not employed, it is suggested that individual focus and programming were increased. Another possibly important confounding, not mentioned by the authors, is that between school effects and treatment effects, since control and treatment conditions did not occur in the same school but in separate school settings. Although the groups did not differ on a variety of pretest variables, school effects rather than treatment effects may well have mediated some of the observed changes. It is not clear why the control subjects were identified in such a negative fashion to school staff. This may have resulted in negative outcomes for the control subjects, raising the possibility that negative effects for the controls, as well as positive effects for the experimental group, may have contributed to the observed differences. Finally, subject attrition also posed problems: only 12 controls and 10 treatment subjects were.../

were available for achievement testing, leaving the possibility of a biased comparison. Despite these design difficulties, the reported impact on adolescents with severe behaviour problems is encouraging.

The cohort of students in this study, which was conducted in 1973-74, was followed up in 1978-79, along with a further 1974-75 cohort, yielding a total of 43 programme and 55 control students in the follow-up population (Safer et al., 1981). School records were available for 89% of the total group. As in the original report, the treatment group maintained their superiority on rates of withdrawal and suspension for disciplinary reasons. Significantly more of the treatment group obtained passing grades than did controls. On a yearly composite achievement measure, the only significant differences, in the last two years, favoured the control group but this comparison is again seriously marred by a high attrition rate.

Rollins et al. (1974) trained sixteen inner city public school teachers in the use of behaviour management procedures. In addition to social reinforcement the teachers relied heavily on a token system with a variety of inexpensive back-up reinforcers (games, records, puzzles, etc) available in a special activity room. The project involved 730 pupils in all, covering an age range of 7 to 14 years. The experimental group consisted of 355 pupils in 16 classes with 375 controls in 14 classes. Assignment to conditions was done randomly. Four schools were involved, with control classes at the elementary age level (up to 11 years) being in a separate school not involved in the treatment process. This arrangement, as in the Heaton et al. (1976) study, creates a possible school/treatment confounding. Control classes for older children were housed in the same schools as experimental classes. About half the experimental subjects received the programme for two school years, the remainder for one. Control classes were 'conducted in a traditional manner, with a single teacher managing each class in a lecture format'. The absence of an equally attractive activity centre in the..../

the control classes is a weakness in the experimental design, so that observed differences could be attributed to non-reinforcement variables.

It is reported that the project group significantly bettered controls in terms of increasing task involvement and reducing disruption. However, it is to be noted that both these comparisons were conducted by analyses of variance, rather than covariance, yet it is evident from the graphed data that initial levels on both measures are quite different, and trends throughout the experimental period are less than obvious. More convincing evidence, based on covariance analyses, is provided on measures of IQ (California Test of Mental Maturity), reading and arithmetic. The project group as a whole, and secondary school age groups within it, bettered controls on the IQ and reading measures, but not on arithmetic. Along with these encouraging results, the authors report increased teacher morale and improved school-parent relations in a difficult inner city setting. It is therefore discouraging to note that one year after the researchers' support was withdrawn, it was discovered that 'the entire programme had been discontinued. Not only were the formal aspects of the programme no longer in use, but in-class observations indicated that teachers had reverted to a more traditional style of teaching with concomitant increase in student disruptions and a reduction in student involvement' (Rollins & Thompson, 1978).

A somewhat more encouraging view of the aftermath of token reinforcement programmes is presented by Dickinson (1974). He followed up 50 pupils who had participated in such a programme while in the 5th and 6th grades (11-12 years old), comparing them with 218 pupils in nearby schools who had not been in reinforcement programmes. Data were obtained at the end of the 8th grade, two years after the token system had ended. The special programme involved the use of tickets as tokens, which could be redeemed for a variety of privileges and toys. They were earned for attending to work, completing assignments and behaviours such as participating in discussions. Reading scores on the Metropolitan Achievement Test at the end of the 6th grade..../

grade were used as a covariate when reading scores at the end of the 8th grade were compared. Students from the reinforcement programme had made significantly greater gains at this 2 year follow-up than the group which had received regular instruction, and they also had fewer students sent to the principal's office for disciplinary problems. It should be noted that school and treatment effects are confounded in the study design. In addition, only those children with complete achievement data and enrolment in junior high schools were included in the analysis. Thus, the 50 target children were selected from an original pool of 94 on this basis. Such a selective analysis leaves the possibility of bias towards the best achievers.

Wodarski & Filipczak (1982) report the outcomes for the Preparation through Responsive Educational Programs (PREP) which aimed to achieve a number of short-term and long-term goals with pre-delinquent adolescents. The principal immediate goals were to expand students' academic and social skills in order to permit them to function more appropriately within their school environments. Thirty experimental and thirty control students were selected 'on the basis of strong evidence of academic or social problems during the prior year'. Allocation to groups was done randomly. The average age was 13 years. PREP operated as a 1-year intervention for all experimental students. The programme had three components: (1) academic training in reading, English and mathematics provided daily in a skills centre during English and Maths period. (During other school periods, instruction was given by regular teachers within regular sessions); (2) social or interpersonal skills training that facilitated immediate and generalisable social skills for problems inside and outside of school; and (3) family skills training that promoted increased involvement of parents in school activities and management programmes in the home. Student rewards were based on successful academic work and social behaviour throughout the entire school day. A behaviour record card was carried by students into all non-PREP classes for rating by their teacher. Rewards included praise, grades, tangible items and activity options such as field trips, films and extra games time. Control students..../

students attended the same school with no alteration in curriculum.

At the end of treatment, there were significant differences in favour of the experimental group on measures of reading comprehension, language skills, mathematics computation and application, number of disciplinary referrals and class grades. At a one-year follow-up, the experimental group had a lesser number of suspensions and higher school attendance than the control, and also performed better on disciplinary referrals and class grades for English and Mathematics. It is not clear, however, whether these differences were statistically significant.

A follow-up some four years after treatment termination was then attempted. Major problems were experienced both in the use of measures of evaluation at this stage and with subject attrition. Measures identical to those originally employed were not available, 'bureaucratic constraints' created problems for approval of certain protocols and information required from the police and employers was not available. Follow-up measures were restricted to self-report, and had no baseline equivalents. Only 40 of the original 60 subjects were available (21 experimental, 19 control). However, comparison of drop outs with available subjects on baseline measures suggested that participants in the follow-up were representative of the total initial sample.

Few differences emerged in the comparison. Three items favoured the experimental group: happier home environment, participation in fewer gang fights, and greater avoidance of trouble. Control subjects, on the other hand, reported reading more, experiencing more guilt about engaging in anti-social activities and having less tendency to engage in aggressive behaviour to secure reinforcers. There was therefore no evidence of long-term maintenance of behavioural changes. Given that 144 items were analysed, the few significant differences can be relegated to chance findings.

Apart from these difficulties with the design of the long-term follow-up,.../

up, other problems are apparent which require consideration in evaluation of the shorter-term results. First, it is assumed that with random assignment to control and experimental groups, pre-test data should not be significantly different and post-test comparisons can justifiably be made. This equivalence is not tested, so the possibility that the two groups differed on initial levels was not excluded. Second, the factors responsible for observed changes in the experimental group are not clear. The design of the study leaves open the possibility that changes may have been due to the extra attention provided rather than to the specific intervention procedures.

Two recent studies (Jason & Ferone, 1978; Jason et al., 1979) although conducted with children of primary school age, are of relevance in attempting to compare different modes of treatment which were somewhat similar to two of the approaches in the present study. They are therefore included in this section. In Jason & Ferrone's (1978) study, teachers experiencing difficulties in managing disruptive, acting-out children in two classes of 7 year olds were provided either behavioural or process consultation. The behavioural intervention included discussions of behavioural principles, feedback concerning contingent praise and individualised interventions. These took the form of positive or negative reports to parents depending on the child's behaviour in class, and retention in class during play breaks contingent on excessive misbehaviour. In contrast, the process consultant used clarifying, supportive and reflective responses to help the teacher better understand classroom difficulties and enhance her ability to work with problem children. During the 7 weeks of consultation and at a 4 week follow-up, problem behaviours were significantly reduced only in the class with behavioural consultation. This finding was based on direct observation of behaviour in class, but both teachers rated the children as being less disruptive following the programme.

Findings from this study need to be interpreted with caution, given a ..../

a number of methodological problems. First, only one teacher was involved in each intervention so 'therapist' and treatment are confounded. Second, there was a significant difference in initial levels of problem behaviour, with the behavioural group worse, so differential regression may have affected the outcome. Third, sample sizes were very small - three children in the behavioural group, four in the process group. Finally, ratings were provided by the teachers directly involved in the interventions, possibly creating bias.

In a similar intervention, Jason et al. (1979) again compared behavioural and process consultation but added two further groups, an ecological consultation group, and a no-treatment control. These were all in different schools. The behavioural and process conditions were essentially the same as in the previous study, with the exception that the former supplemented teacher attention with the earning of points exchangeable for small toys. The ecological consultation consisted of discussion of actual and ideal classroom environments, and restructuring class groupings and seating arrangements. There were nine children, again with disruptive, acting-out problems in both ecological and behavioural conditions and seven in process and control conditions. Attrition problems, however, left only 3 to 5 children in each classroom. They were aged 6 to 9 years. Consultation lasted two months and a one month follow-up followed the intervention.

Significant reductions in observed and rated behavioural problems occurred only in classes which were provided with behavioural consultation. Observed improvements in behaviour persisted at follow-up. Children in the process condition showed significantly worse problem behaviour both during treatment and at follow-up. Significant increases in reading and arithmetic were found in the process condition, while children in the behavioural condition showed decreased reading scores. These apparently contradictory behaviour-achievement relationships are possibly attributable to a combination of small sample sizes, with changes by individual children having excessive influence,...../

influence, and the uncontrolled pre-treatment differences on both behavioural and achievement measures. As with the previous study, confounding is present, this time involving treatment/therapist and treatment/school. The authors also comment on the unreliability of some of their measures and lack of long-term follow-up.

In summary, these two studies represent important attempts to compare behavioural approaches with other treatment modalities, but, unfortunately, in view of their major methodological shortcomings, little confidence can be placed in their results.

(f) Concluding note

To conclude this section, there is evidence that a number of methods are successful in modifying problems of behaviour and deportment in the classroom. Methods involving token procedures may be more powerful than those limited to alterations of teacher attention, but their use may be more controversial. From the perspective of the present study, an important deficit in the literature, at the time of planning, was the relative shortage of studies involving secondary school children. This was particularly true of investigations of social reinforcement - one of the primary concerns of this study - and it is interesting that the literature subsequently has not filled the gap. A further shortcoming was that most studies, in addition to employing young children as subjects, were conducted in experimental classes or specially composed 'adjustment' classes of problem children, somewhat different from regular school settings, and with limited numbers of children. The early studies, therefore, left many questions relating to generalisability of findings, some of which the present study was concerned to pursue. More recently, larger scale interventions in ordinary secondary school settings have been reported. Some of these have incorporated designs with control/comparison groups, as in the present study. It is difficult, however, to draw firm conclusions from these studies, given the presence of.../

of design problems in each one of them. One encouraging feature of most of these controlled studies is their use of multiple measures of change.

Issues of class management and control should ideally be viewed as a means to an end and not as an end in themselves. The objectives of fostering and encouraging children's learning should be paramount and it is important to ask what contributions behaviour modification can make here. It is to these questions that we now turn.

(ii) Interventions for academic behaviours

(a) Introduction

The present study included measures of reading comprehension, and of verbal and non-verbal ability in the assessment battery as a means of gauging what gains, if any, might be associated with increases in task-attention, or task-relevant behaviour. The question of academic gains was not a primary focus in the methods employed, in that reinforcement was not arranged to be contingent principally on factors such as academic output, rate, quality of product and so on, but rather on attention to task and task requirements. However, it would seem appropriate for any approach to classroom management to be concerned with implications for achievement. This section will examine relevant areas of the behaviour modification literature dealing with academic outcomes. To begin with, some issues to do with the rationale for intervening with academic behaviour will be considered, along with questions about the most appropriate focus for such interventions.

If behaviour modification is construed as a process of learning, unlearning and relearning, then it is not surprising that academic subject matter has become an appropriate area of concern. Staats (1968) was one of the first psychologists to conceptualise academic activities as operant responses. With reading, for example, the process can be regarded as one in which the child makes a response to a written stimulus. The teacher's praise or approval following the response can function as a reinforcer, and can be used in

a..../

a systematic fashion to assist the learning process. Extending the notion of reinforcement, Staats (1973) went on to argue that schools should be encouraged to mobilise their ample stock of reinforcers in order to bring them to bear on children's learning and to enhance their motivation. The studies reviewed in this section show the benefits of reinforcement procedures brought to bear on such responses such as rate of output, number of items completed, number correct and so on. The impact of such procedures can no doubt be increased when they are integrated with a thorough task-analysis of particular academic functions.

In ideal circumstances, learning carries with it its own natural reinforcement. Reading, for example, becomes 'its own reward' as a pleasurable activity, one that yields information and knowledge, and so on. This 'natural' reinforcement is logically related to the activity of reading. Ferster (1967) drew a distinction between such forms of reinforcement and 'arbitrary' reinforcers which bear no logical relationship to the behaviour. Words of praise from an adult, a token, or a score on a test, are equally arbitrary in this sense, in relation to a reading response. However, such arbitrary reinforcers may be needed in the early stages of learning before a child engages in reading 'for its own sake' and before the more natural reinforcers take over. It is in this sense that arbitrary reinforcers may have to be used with children who experience difficulty and for whom particular learning tasks may have become an aversive activity. As Staats (1973) pointed out:

'...the time spent in the classroom by problem children is almost entirely a waste. They have a very low rate of learning responses. They do not attend to what is being said in class, they do not read, they do not work problems.... The reason that they do not learn is that they have very few learning trials.' (p.222)

What is being suggested here is that 'problem' children do not learn adequately because, for much of the time, they are not actively engaged with the task they have been set. The assumption is implicit in much behavioural work that increasing task attention and engagement is likely to result in academic gains. This assumption, however, has been strongly challenged. Much..../

Much discussion has been directed at the question of whether it is more effective to modify social behaviours such as task attention, compliance, listening to instructions, etc., with the expectation that this will generate academic gains, or whether it is more useful to reinforce academic products directly. The former strategy has been criticised by Winett & Winkler (1972) who questioned the 'rigid preoccupation with order and control....where children are required to be still, to be silent, and to obey' (p.499). They challenged the motivation behind this strategy, suggesting that rather than contributing to useful learning, it most often leads to a strengthening of the status quo, and to classrooms excessively concerned with control. Justification for the approach usually centres around the assumption that in order for children to improve academically they must first be taught to pay attention. Disruptive behaviour is incompatible with attending and logically should be decreased prior to attempting to increase the amount of effective learning. The evidence for this position, however, and the opposing one which favours direct reinforcement of academic products, is inconclusive.

A number of workers have documented the extent to which academic achievement and social behaviours are correlated. Thus, behaviours rated by teachers, such as attention, persistence with tasks, compliance with teacher demands and ability to follow directions, have all been found to correlate highly with various achievement measures and teacher grades (Davidson & Greenberg, 1967; Swift & Spivack, 1968, 1969). Longitudinal studies have demonstrated the utility of teacher ratings of 'attention' and 'co-operation' in predicting later achievement (Meyers et al., 1968) and the extent to which aggressive and disruptive behaviours can forecast academic failure (Feldhusen et al., 1970). It is a rather common-sensical notion that the more time that a pupil spends on a task, the more he should learn. Harnischfeger & Wiley (1975), indeed, suggest that the total amount of active learning time on a particular instructional topic is <u>the</u> most important determinant of pupil achievement on that topic. There is some evidence for this position both when learning time.../

time is construed in gross terms (allocated time, or quantity of schooling) or in quite focussed terms (actual engaged time). In the first category, Stallings (1975) found in a study of American schools that the length of school day varied as much as two hours per day between schools. The length of school day was one of the variables strongly correlated with achievement in both reading and mathematics in an evaluation of 150 Follow-Through classes. Wiley & Harnischfeger (1974) related the average number of hours of schooling pupils received, derived from average daily attendance figures, length of school day and school year, to achievement in verbal ability, reading comprehension and mathematics and found clear, positive relationships. Absenteeism is one obvious way in which allocated time is reduced and predictably, it tends to be associated with lowered achievement. Fogelman (1978), using data from the National Child Development study to examine the relationships between school attendance at ages 7 and 15 and achievement at 16, observed that children with high attendance levels obtained on average higher scores on tests of reading, comprehension and maths. These relationships did not differ by social class.

The allocation of time to a learning activity, however, or having a pupil physically available, does not ensure that time will necessarily be spent on-task or engaged in academic activities. The opportunity to make use of available time will be affected by factors such as disruptions, distractions, lack of interest and poor persistence. 'Learning time' therefore requires a more precise definition, in terms of actual engaged time. With this more refined approach, Fisher et al. (1977), observing a sample of six-year-olds over an eight week period, found strong relationships between active learning time and achievement for both reading and maths. They also related both allocated and engaged time to achievement and found that the latter relationship was stronger. Arlin & Roth (1978) carried the analysis of engaged time even further in the context of children's reading, in pointing out that while children may apparently be 'on task' in the sense of having a book out and looking at it, they may not actually be <u>reading</u> it. They therefore distinguished between 'time-on-task' and 'time-on-reading' with forty-two 9-year-olds and found.../

found the latter to be more strongly associated with reading gains. The link between work activity/task-attention and achievement has been confirmed in several other studies, with a range of ages. Cobb (1972) with 11-year-olds found that the best predictors of achievement were 'attending' and 'task-orientated conversation' amongst eight categories selected for observation. Lahaderne (1968) reported correlations of .39 and .51 between the frequency of attending behaviour and the achievement of 12-year-olds. Samuels & Turnure (1974) found similar correspondence with a 6-year-old sample. From these findings, negative consequences for achievement would be predicted for inattentiveness, as shown by McDonald (1975). Relatedly, Evertson (1980) found that low-achieving junior high pupils were engaged on task 40% of the time in academic activities compared with 85% for their high-achieving colleagues. Low-achievers also had more 'dead time' in which nothing happened.

On the basis of such findings, it would seem to follow that management procedures which could increase on-task rates could have important implications for academic achievement, as wellas serving to reduce undesirable behaviour. Cobb and his colleagues (Cobb, 1972; Hops & Cobb, 1973) at the University of Oregon have pursued this hypothesis. Having demonstrated the relevance of specific behaviours, such as 'compliance', 'looking around' (negatively related), 'attending' and 'volunteering' to achievement in observationally based studies, they argued that these are 'not academic behaviours per se, but, rather, the first components in a chain of correct academic responding.' (Hops & Cobb, 1973). These behaviours are viewed as 'survival skills', necessary, but not sufficient for successful academic funcitoning. If these survival skills can be improved and strengthened, academic gains will follow. This hypothesis has been supported for reading achievement (Cobb & Hops, 1973; Hops & Cobb, 1973; Greenwood et al., 1977) and for both reading and maths (Walker & Hops, 1976) with improvement being measured on standardised achievement tests. Evidence also exists, however, to show that increasing attentive or reducing disruptive behaviour does not necessarily have an effect on academic performance..../

performance. Ferritor et al (1972) found that when tokens were made contingent upon attending behaviour, disruptive behaviour decreased, but academic performance was not affected. Only when tokens were made contingent upon both attending and academic achievement did performance improve. Harris & Sherman (1974) reported similar effects. A more parsimonious approach may therefore be one in which academic products are directly reinforced, since this may not only increase academic output but also improve levels of attention and decrease disruptive behaviour. Ayllon & Roberts (1974) for example, provided tokens to disruptive 11-year-olds for completing reading assignments correctly. When academic performance increased, disruptive behaviour decreased, although no specific contingencies were designed for disruptive behaviour. Hay et al (1977) selected two groups of 5 boys from among ten normal 8 to 10 year olds. Each group was exposed to two treatments counterbalanced for sequence. Academic contingencies were applied first in one group and on-task contingencies first in the second group. Teachers delivered a fixed amount of praise contingent upon either on-task or arithmetic/reading performance depending upon the phase of the study. On-task contingencies improved on-task levels only, while product contingencies improved not only on-task levels, but rate and accuracy of output as well. The wider implications of a focus on academic products have similarly been demonstrated by Ayllon et al (1972) and Winett & Roach (1973). This outcome has not always been obtained, however. Walker & Hops (1976), in a comparison of reinforcement of academic performance, reinforcement of 'facilitative nonacademic' responses and the use of these procedures in combination, found no differences between the three strategies, although all three improved academic performance.

Drawing firm conclusions in this debate is hampered by the tendency of some contributors not to cite or consider evidence contrary to their own position (e.g. Ayllon & Rosenbaum, 1977; Lahey et al., 1978). It is possible that studies which have not shown positive correlations between increased on-task behaviour and improved academic performance have adopted 'on-task' definitions which.../

which have been too broad, leading to a focus on behaviours only tenuously if at all, related to actual academic performance. The studies on 'survival skills' at Oregon, on the other hand, have grown out of a series of investigations of which social behaviours facilitate or enhance performance and their focus is accordingly more refined. The need for a sharper focus is also highlighted by those workers who distinguish between 'allocated time' and 'engaged time', as discussed earlier. It is also apparent that many of the studies cited here which failed to find achievement spin-offs from focussing on social behaviour have based their findings on assessments in the short term (as do most studies in this area). The effects of significantly improving social behaviours considered conducive to learning on achievement in the longer term have not received thorough investigation. One thing that is clear is that targets for behaviour modification must be selected with care, and with consideration of the potential benefits for the parties concerned - both children and teachers. It is also incumbent on workers to base such a selection on demonstrated linkages between target behaviours rather than on suppositions of such relationships.

In conclusion, the debate on the most appropriate focus for reinforcement was just beginning at the time of planning the present study. At that time, the emphasis on reinforcement of social behaviour rather than academic products was more typical, and this was the orientation adopted in this study. Although arguments for and against such a stance have been marshalled in more recent years, the debate is still not resolved.

(b) Studies emphasising social reinforcement

Compared with the number of studies demonstrating the impact of teacher attention on social behaviour, the number documenting its effectiveness with academic achievement are few, with token procedures being more commonly employed. However, Kirby & Shields (1972) showed that the rate of arithmetic problem-solving could be increased by praise and correctness feedback, albeit with only one subject, a 13-year-old boy. Stromer (1975) employed praise along with correctness feedback and modelling to successfully modify letter and number..../

number reversal difficulties with seven children aged 6-8 years in special education classes. Hasazi & Hasazi (1972) reported success in correcting habitual digit reversal in an 8-year-old boy through the use of differential attention. Clark & Walberg (1968) employed an untreated control group in their experimental design in working with 110 potential school drop outs in an after-school remedial programme. These children were from 10 to 13 years of age and from one to four years behind in their school work. They were assigned randomly to nine classes with 62 children in the experimental group and 48 children in the control group. Teacher praise was made available to both experimental and control subjects, with children making a tally mark on a card for each praise comment received. The description of procedures is rather inadequate, in that it is not clear at what praise was being directed. After three weeks, the teachers of the experimental group were asked to double or triple the number of rewarding comments, while control teachers were asked to 'keep up the good work'. No data on praise rates are presented. At the end of the second three week period, the children were administered the SRA Reading Test. The experimental group scored significantly higher on this measure than did controls, with IQ being employed as a control variable in a covariance analysis. The results are, however, rendered uninterpretable by the somewhat curious omission of a pre-test on the reading measures, with a post-test only being examined.

The relationship between teacher praise and academic achievement has also been studied extensively in the field of classroom observational research which has developed quite separately from behaviour modification investigations. Reviews of this work have produced inconsistent findings. Rosenshine (1970) and Dunkin & Biddle (1974) noted that most studies had shown no clear relationship between praise and achievement, criticism was usually negatively related to achievement, and that acknowledgement or acceptance of responses was usually positively related to achievement. Stallings & Kaskowitz (1974), on the other hand, found that both praise and corrective negative feedback showed positive relationships with reading and mathematics achievement. Variation in coding conventions make interpretations across..../

across these studies rather difficult. An additional difficulty, from the perspective of behaviour modification, is that many of these studies do not place any special emphasis on the contingency between child response (e.g. correct reading) and teacher feedback (e.g. praise and approval). It is therefore possible that, since the kinds of close connections which would be sought in a behaviour modification study are not necessarily highlighted in these studies, their findings are of questionable relevance within a behavioural framework.

Assuming that positive reinforcement can raise scores on academic measures, questions can be raised as to how such improvements should be interpreted, as Ross (1976) points out. For example, it has been shown that reading comprehension scores can be increased by reinforcement of correct answers in the absence of any training in the processes involved (e.g. Lahey et al., 1973). This leads Ross to suggest that in these circumstances one may not have increased comprehension at all but merely the motivation of the child to emit responses, the capacity for which had been present all along. While distinctions between gains in 'real ability' and increased motivation are perhaps difficult to make, the point made by Ross is an important one, and one which has tended to be ignored by investigators in this area. It is nevertheless important to consider that even if 'real' ability has not been materially affected, the gains associated with reinforcement procedures may have much personal significance for children who have been accustomed to failure and who begin to respond with increased drive and confidence. An improved level of functioning can, of course, have wide implications for self esteem.

Interpretations of improved scores in terms of motivational factors can also be made where the measures taken are of ability or intelligence levels rather than achievement levels. In addition to gains that may be motivationally based, changes on these tests may also be a function of the extent to which their content reflects the academic repertoires which are directly reinforced in the classroom.

(c) Studies emphasising token reinforcement

Most classroom studies which focus on academic behaviour have employed some form of token reinforcement system. Early studies by Staats and his colleagues (Staats et al., 1964) paved the way by developing a technology for the experimental analysis of reading behaviour. An early venture into the classroom by Wolf et al. (1968) showed the effects of token procedures in an after-school remedial programme for low achieving 11-12 year olds who were at least two years below the norm for reading. Children were allocated in blocks matched on reading scores to the experimental programme or a control condition with 15 subjects in each. In the experimental programme, tokens, which were exchangeable for sweets, novelties, field trips and other items, were delivered for correctly completing classroom assignments. Controls received standard remedial procedures. At the end of the one-year programme, experimental subjects showed significantly higher achievement test performance (Stanford Achievement Test) and better school grades than controls. The experimental subjects gained 1.5 years on the achievement test while controls gained 0.8 of a year in the same time. Similar impressive gains have been reported in remedial settings by Clark et al., (1969); Kaufman & O'Leary (1972) and Pelham (1974). Interpretation of these studies however, is complicated by the inability to separate the effects of token reinforcement from the affects of attendance in a remedial group per se.

Hewett, Taylor & Artuso (1969) rewarded a group of emotionally disturbed 8-11 year-old children in a specially structured experimental classroom with checkmarks (leading to tangible rewards like sweets, prizes and extra time in arts and crafts) for the classroom behaviours of being on time, following directions and correctly completing assignments. A control classroom condition consisted of 'any classroom strategy' other than the use of checkmarks and tangible rewards. Significantly better progress in mathematics was found for the experimental group, but the token programme did not influence reading and spelling gains.

The age spread in applicability of token procedures in the area of academic

performance...../

performance is impressive. They have been applied successfully with children beginning kindergarten in the Follow-Through programme for disadvantaged children (Bushell, 1974) through secondary school (Glynn, 1970; Chadwick & Day, 1971; Kirby & Shields, 1972) to college level (Lloyd & Knutsen, 1969). The main focus with school children has been on basic achievement in functions such as reading, arithmetic and spelling. It is possible that functions such as these are easily quantified and thus readily satisfy the need for objective data in behavioural studies. However, other skills have also been improved. Thus, gains in quality of handwriting (Robin et al., 1975), writing of compositions (Maloney & Hopkins, 1973) and increases in creativity in short story writing (LaGreca & Santogrossi, 1975) have been reported - all in the context of token programmes.

(d) Studies employing control or comparison group designs

The studies which require consideration in this section have already been described in some detail in Section (i) of this chapter. Briefly, Heaton et al. (1976) found significant differences in favour of their behavioural group on the reading section of the Wide Range Achievement Test (WRAT), but no differences on arithmetic and spelling. In the reading section of this test, children orally read individual words. The four year follow-up to this study focussed on composite (oral reading plus comprehension) rather than individual WRAT subtest scores, and the only significant differences in the last two years of follow-up favoured the controls. Rollins et al. (1974) reported significant gains for their behavioural treatment group on the California Test of Mental Maturity (yielding an IQ score) and on the reading, but not the arithmetic subtest, of the California Achievement Test. The reading test here yields a composite score for the reading of single words and for comprehensio Dickinson (1974) similarly found significant gains on a composite reading score on the Metropolitan Achievement Test at a two year follow-up for children who had participated in token reinforcement programmes. Finally, Wodarski & Filipczak (1982) reported end-of-treatment differences in favour of their experimental.../

experimental group on measures of reading comprehension, language skills and mathematics computation and application. These measures were not available at follow-up.

These four studies are the only ones available which examined behavioural applications with a control group design within the regular secondary or junior high school context. They all employed token reinforcement in one form or another. While the spread of effects across achievement and ability measures is encouraging, it should be recalled that none of these studies is free of design problems (e.g. attrition, differential regression, confounding of treatment and school effects). The interpretation of findings therefore has to be treated with caution. In addition, only the study by Dickinson reported positive follow-up data. Criticisms specific to each study have already been considered.

A highly critical view of behavioural research in the area of reading achievement - which figures prominently in the above studies - has been offered by Lahey (1977). He adopts the position that reading consists of obtaining meaning from the content of material read: it is not merely the pronouncing of individual words. Assessment of reading comprehension is therefore for Lahey the most important criterion in evaluating reading performance. Of the 36 studies of reading he was able to identify (covering primary- and secondary-school age children in various educational settings), he found that this body of research had relied too heavily on either limited or inappropriate measures of reading. Ten studies report data only on the reading of single words, two on the oral reading of sentences, five on rate of progress through programmed workbooks, and nine report results as a 'composite' reading score in which scores for reading single words cannot be separated from scores based on measures of comprehension. Only nine studies reported any sort of independent measure of comprehension. Results of these latter studies tend to be less favourable than the others. Only three studies reported positive findings for reading comprehension, and, of these, two used inadequate designs lacking control groups..../

groups and only one had an acceptable between-groups design and appropriate standardised measures. Lahey concludes that:

'Because the behaviours that are referred to as "reading comprehension" are widely considered to be the most important reading behaviours, this analysis of the research literature seriously undermines the conclusion that positive reinforcement can produce increased reading achievement." (p.30)

Re-examining the four controlled studies above, only one (Wodarski et al.) employed what would be in Lahey's terms an acceptable measure of reading comprehension. His critique is of special relevance to the present study in which reading ability was assessed. The measure employed was a test in which the child supplies a missing word, from several possibilities given, which would meaningfully complete a sentence.

(e) Concluding note

To summarise this section, it is evident that although many studies have focussed on the modification of academic achievement and have claimed positive results, a variety of design problems hamper firm conclusions. The few available studies employing control group designs are subject to important methodological limitations. All of these employed token reinforcement in some form, rather than emphasising social reinforcement, as planned in the present study. Where gains in academic performance have been claimed, it is noted that a question of interpretation exists, as to whether the improvement reflects 'real' gains or increased motivation. In measuring academic performance, it is important that meaningful measures are selected. Of special relevance for the present study is the recommendation that where reading ability is concerned, measures of reading comprehension are the most appropriate. A wider debate has arisen concerning the question of whether one should focus on social behaviour directly or on academic products when one is concerned with issues of classroom management and the fostering of achievement. At the time of planning, the former focus was in favour, and the study was designed with this orientation. Subsequent studies on engaged time would seem to support this choice but the debate is by no means resolved. Having considered classroom management and academic achievement issues, we now turn to a third sphere of importance in children's school..../

school lives - that of interpersonal relationships.

(iii) Interventions for problems of social interaction

(a) Introduction

In addition to coping with academic demands and the requirements for behavioural control in their school life, children also have to acquire the skills which make for effective interpersonal relationships in their peer group. Failures in this sphere may have important implications for their later functioning (Combs & Slaby, 1977; Macmillan et al., 1978) especially if peer difficulties take the form of aggressive-disruptive behaviour (Conger & Keane, 1981). These deficits may also have important implications for classroom functioning. Thus, some teachers may selectively ignore students who are perceived to be socially and academically inferior (Brophy & Good, 1970), they may direct a disproportionate amount of sarcasm and threats at them (Khlief, 1976) and direct few nonverbal signs of approval or liking towards those who are socially isolated (Lyon, 1977). Intervention with these problems is therefore of importance in attempting to alleviate current difficulties and to forestall complications in the future. An important concern in the present study was that of helping children who seemed to be experiencing interpersonal difficulties, as indicated by sociometric measures. This section will therefore examine relevant literature in this area. It will be concerned primarily with studies involving teacher-administered reinforcement and those methods which seem compatible with regular classrooms which are not geared to conveying intensive help to such children. In addition, it will of necessity be concerned with young children, since there are no examples available of behavioural applications in secondary school settings, which is a major gap in the literature.

(b) Procedures involving reinforcement by teachers

Studies of the application of operant procedures to shaping children's social behaviour have tended to focus on the young child. In two early studies of social interaction, it was found that two pre-school children who spent little time interacting with peers attracted or maintained adult attention when engaged in isolate play (Allen et al., 1964; Johnston et al., 1964). Teachers initially shaped approximations to peer contact by attending when the children merely approached..../

approached and watched peers play and later attended only to actual peer intervention. This contingent-reinforcement procedure resulted in a marked decline in isolate play and a two-to-threefold increase in social play. During the reversal phase, teacher attention to solitary play resulted in a return to baseline levels of such play, while reinstatement of reinforcement restored increased peer interaction. Follow-up 26 days after the completion of the study (Allen et al., 1964) and during the following school year (Johnston et al., 1964)showed good maintenance of these gains. However, the quick return to baseline levels of social interaction during reversal phases suggests that these gains were dependent on the continuation of adult reinforcement. As pointed out by reviewers of intervention studies in this area (Combs & Slaby, 1977; Wanlass & Prinz, 1982) follow-up is either neglected or shows poor results. For example, O'Connor (1972), in a study comparing the effects of modelling and reinforcement procedures, increased social interaction in isolate pre-school children through reinforcement alone, but found a reversal to baseline levels three weeks after treatment.

Hart et. al. (1968) shaped co-operative interactions with a 5-year-old girl by means of contingent social reinforcement. She was initially reinforced for verbalisation when in the proximity of peers, then for participation in potentially co-operative situations and finally only for actual co-operative play. Reliable increases in co-operative play resulted, in contrast with the lack of impact of reinforcement delivered noncontingently. The dependence on immediate adult reinforcement was again demonstrated across contingent/non-contingent phases.

Another form of interpersonal difficulty - aggressive behaviour - has also been dealt with by adult reinforcement procedures. Brown & Elliot (1965) and Pinkston et al. (1973) employed a combination of ignoring inappropriate and aggressive behaviour and reinforcing non-aggressive or co-operative interactions. 'Time-out' procedures, involving brief isolation, have also been used to manage aggressive behaviour while co-operative behaviour was simultaneously shaped (Allen et al., 1972). Disruptive and aggressive social behaviour has also been modified in the classroom by means of token reinforcement, as described in an earlier section (e.g. O'Leary & Becker, 1967; O'Leary et al., 1969).

(C)...../

(C) Problems, limitations and alternatives

It is evident in surveying the available literature on procedures applied by teachers that there are a number of shortcomings, especially from the perspective of the present study, where guidance was sought for potential applications in secondary school settings. The various gaps in the literature are well-illustrated in a review by Wanlass and Prinz (1982). Of the thirteen studies employing adult-mediated procedures which they identify, none involved children of secondary school age in regular settings. There are indeed no such studies amongst the 42 identified in the entire review, emphasising the neglect of the interpersonal problems of these older children. The generalisability of findings from studies with children at pre-school or primary school level is quite unclear. McNamara & Harrop (1979) suggest that teacher praise and attention may be less reinforcing with older children because the pupil may view it as alienating him from the peer group or because the peer group status achieved by 'bucking the system' may be greater than any available teacher-controlled reinforcer or sanction. Whether these limitations actually hold for interventions with older children who are experiencing peer-interaction difficulties is a question that has not been researched.

Two other noteworthy issues emerge from Wanlass and Prinz' (1982) review. First, none of the operant studies used sociometric status as an outcome variable. The emphasis instead has been on frequency of interaction, which has been criticised because low interaction rate is not necessarily indicative of maladjustment (Gottman, 1977). Second, only four operant studies conducted follow-up; the longest period being 6 months. In the review as a whole, 22 studies omitted follow-up and a further 11 limited it to a period of one month after treatment or less.

Considering studies of teacher-administered contingencies more generally, other dissatisfactions have been expressed. Thus, adults intervening to provide reinforcement may well interfere with and disrupt the interaction (O'Connor, 1972) and no little skill may be required in ensuring that intervention is appropriate and..../

and sensitive (Roedell et al., 1977). Strain et al. (1976) also suggest that these methods require continuous teacher presence otherwise the effects disappear and, additionally, that these efforts are likely to be inconsistent as children move around a variety of classroom environments. Finally, adult-dependent methods may fail to exploit the important contribution of peers in social interaction and the central role of the cues, feedback and reinforcement that they provide (Strain & Shores, 1977). These authors called for intervention procedures to pay more attention to the reciprocal nature of interpersonal processes.

One study which proceeded in this manner was that of Walker & Hops (1973). Working with 12 isolated children, aged 6-8 years, they showed that interaction could be increased by rewarding (with points exchangeable for back-up reinforcers) both the withdrawn child and classroom peers. They used a film showing models of interaction as a training method and examined three experimental conditions:- (1) rewarding a withdrawn subject for peer initiations, (2) rewarding trained peers for initiation by a withdrawn subject and (3) rewarding the withdrawn subject for peer initiations and also peers for subject initiations. While all conditions increased the withdrawn subjects' interactions, the third condition was the most successful. In an extension of this work, focussing on different components in interaction, Walker et al. (1979) found with eighteen 6-12 year-olds that by simply reinforcing social initiations or responses to initiations, the frequency of these behaviours could very readily be increased but at the expense of making interactions very brief and artificial. Reinforcing continuing social interaction over time resulted in more meaningful exchanges.

In a further criticism of adult-dependent procedures, it has been pointed out by Combs & Slaby (1977) that these operant methods in general depend on the occurrence of the desired behaviour, or some approximation to it, which can then be "shaped" up to the required form. If relied on exclusively, such methods may be excessively time consuming, and with extremes of isolated or aggressive.../

agressive behaviour, they may be quite unhelpful. These authors also point to the limitations of this approach where specific new or complex social skills need to be taught, such as the skills involved in initiating interaction or solving social conflicts. Unfortunately, the above methods are perhaps the ones that best 'fit', or are most compatible with, ongoing teaching activities, and the alternatives call for more structured and explicit provision for helping the children in need. These include film and live modelling, coaching-instructional techniques, group methods and problem-solving approaches, and are comprehensively reviewed by Conger & Keane (1981). These approaches will be briefly described here.

O'Connor (1969) demonstrated dramatic changes in the level of social interaction of an experimental group of six pre-school isolate children following exposure to a film in which an initially withdrawn child engaged in increasingly complex social interaction. However, there was no follow-up and the bulk of the effect seemed due mainly to two of the six children. In a subsequent study with 31 pre-schoolers (O'Connor, 1972) the same modelling method, and modelling combined with shaping, proved more successful than shaping alone, and the effect persisted at a 9-week follow-up. While Evers & Schwartz (1973) successfully replicated the effects of the O'Connor film - although with no control group - Gottman (1977) failed to do so. Furman et al. (1979) obtained mixed results with live modelling procedures in which 24 withdrawn pre-schoolers were assigned to either ten socialisation sessions with a younger child, or with a same-aged child, or to no treatment. Both treatment groups benefitted but only the younger-age peer condition differed from controls.

An extension of the basic modelling paradigm has been developed by Gottman et al. (1976) to include coaching by an adult on making friends and interacting positively with peers, and how to take the perspective of others. A nine week follow-up showed significant changes in sociometric position but there were only two nursery school children in the treatment group.

Allen et al. (1975) undertook group treatment of twenty-three 10-11 year old sociometric isolates, meeting some 13 to 17 times, on a weekly basis, for sessions lasting 50 minutes each. The programme proceeded in three phases - (1) simple interactive games, (2) games emphasising poise and flexibility within a social group and (3) social play in the playground outside the group - and children were reinforced with tokens for appropriate behaviour. Sociometric measures at end of treatment and at follow-up five months later showed significant gains for the treatment children, while untreated controls showed only minimal changes. The follow-up, however, was marred by attrition problems which render this difference unreliable.

The final approach to be considered here, that of interpersonal problem solving, is exemplified by the work of Spivack and his colleagues (Spivack & Shure, 1974; Spivack et al., 1976). This is a model which involves the teacher quite centrally but passes much of the responsibility for conflict-resolution over to the children themselves. These investigators have defined and measured a series of interpersonal cognitive problem-solving skills that have been shown to relate to social adjustment at various ages. These skills include: sensitivity to interpersonal problems, tendency to link cause and effect spontaneously, readiness to view possible consequences of actions, ability to generate alternative solutions to problems, and to conceptualise step-by-step means for reaching specific goals, and the ability to view situations from the perspective of other involved individuals. Their programme was developed for classroom use and consists of daily structured activities and discussion, taking 20-30 minutes per day for some 12 weeks. Preliminary teaching involves paying attention to the linguistic and affective concepts presumed to be pre-requisites for effective problem solving. Shure & Spivack (1980) found that training in conceptualising alternative solutions and consequential thinking with 4-5 year-olds was significantly related to improved social adjustment as rated by teachers. A similar kind of programme is being developed for adolescents (Platt et al., 1975) but this has not yet been evaluated.

(d) Concluding note

There is little direct guidance in the above studies for interventions geared to the ordinary secondary school. Most have been conducted with pre-school or very young children so their relevance remains in question. Although the adult-dependent reinforcement methods, which seem the most suited to regular teaching situations, have been criticised as limited in reach and conception, little evidence has been advanced to support this. In school settings where the more extensive and innovative methods, possibly requiring a much broadened conception of the school curriculum, are not feasible, such methods seem worthy of investigation. This was the rationale for their inclusion in the present study as a means of tackling interpersonal problems. Finally, on a methodological note, little attention has been addressed to the issue of maintenance of treatment effects, so provision for follow-up in such studies needs to be made. The behaviours considered so far in this review - inappropriate social behaviour, academic and interpersonal behaviour - are probably the ones which have received most attention in classroom behaviour modification research. The next section will examine the relatively neglected area of assessment of changes in attitude and personality variables and suggest that the inclusion of such measures is also worthwhile.

(iv) Changes in attitude and aspects of personality

(a) Introduction

Much behaviour modification research has been preoccupied with narrow response measures - often relying exclusively on observational data on the specific behaviours being focussed on. The implications of change in one sphere of functioning for other aspects of behaviour have tended to be neglected. This section is concerned with aspects of attitude and personality change in behaviour modification programmes. The present study sought measures in both these areas as a means of monitoring changes that might be associated with improvements in the specific classroom behaviours which were the targets of intervention.

Within the context of a school-based behavioural programme, in which a child, hopefully, experiences a change in teachers' manner of relating to him, with more frequent praise and less criticism, it would seem important to explore possible concomitant changes in the child's attitudes to different aspects of school. Good & Brophy (1974) suggest that the teacher's treatment tells the child what behaviours and achievements the teacher expects from him and this in turn affects the child's self-concept, achievement motivation and level of aspiration. On this perspective, more positively-toned teacher behaviour can have important implications for the child's view of himself and, presumably, the way in which he views the school. Improved academic performance and general coping within the school can also contribute to attitude change.

Attitude measures represent one form of self-report by the child which comprises an important perspective on the process of change within a treatment study. Personality measures can also provide useful self-report data, especially where they allow exploration of the worries, concerns and feelings of anxiety that a child might experience not only within the school context but also in wider living. Improved functioning within the school could be usefully tapped by such measures.

Attitude and personality concepts have been largely ignored by behavioural models which have stressed overt, observable behaviour. However, with the growth of models of cognitive-behavioural intervention (Kendall & Hollon, 1979) and approaches to the understanding of behaviour which stress person-situation interaction(Mischel, 1973), these notions, in one form or another, have received greater recognition. 'Person' variables, whether in the guise of attitudes, beliefs, expectancies, attributions or perceptions are increasingly called upon in the explanation of behaviour within a behaviour modification framework (Mahoney, 1979). This occurs without the implication of static dispositions, such as traits, or associated factors assumed to be 'within the person', which is the stereotype of traditional models of personality.

Within classroom behavioural programmes, as in the wider research context, little attention has been paid to the measurement of possible attitude and personality change. Nevertheless, prominent workers in the field have underlined the relevance of such approaches. Miron (1971) for example, suggested that teachers' praise of children's positive behaviour should encourage the child to believe that he was more competent and independent, thus enhancing his self-concept. O'Leary & O'Leary (1972) similarly suggested that the child's gaining new social or academic skills would lead to more positive reactions from his environment, and this might have the additional benefit of having the child think better of himself. As argued by Good & Brophy (1974) it is likely that teachers are potent figures in helping to determine how children evaluate themselves, especially within the school context, but perhaps more widely as well. Hargreaves (1967), for example, in a study of secondary school boys, suggests that the self-images, attitudes and aspirations of the pupils are inextricably interwoven with the expectations of the teachers for the boys' behaviour and achievement. Attention to attitude change is also encouraged by observations that attitudes to school may be closely related to achievement (Barker-Lunn, 1969; Burns, 1982). Improvements in attitude may have positive implications for this area.

(b) Studies of attitude change

No examples could be found in the literature in which aspects of attitude change (including self-perceptions, or attitude to the self) in the context of behavioural programmes were explored within regular secondary school settings. Gurney (1979) has reviewed intervention studies dealing with self-concept and self esteem with younger children and those in special education settings. He reports a balance between positive and negative findings for attitude change, and suggests that the inclusion of unpublished studies was responsible for this balance. All the published studies produced positive findings, indicating a clear publication bias towards positive results.

In the absence of secondary school applications, two examples of interventions..../

interventions in ordinary primary schools will be described. Buys (1972) worked with nine children, aged 9-11 years, who were identified in three separate classrooms as treatment subjects, on the basis of disruptive behaviour. They were described as difficult to control, with hyperactivity, non-task-oriented behaviour and verbal outbursts the most prevalent problems. Three control children were chosen at random in each of the three classrooms and were not exposed to the treatment variable. The intervention procedure consisted of a combination of public and private praise for appropriate behaviour. Reduction in deviant behaviour in the target children was shown in an ABAB design, over an 11 week period. Control children's behaviour also improved, but only in the first treatment phase. Attitude scales were administered three times: before the baseline, and at the end of each treatment phase.

From baseline to the end of the first treatment phase (after 3 weeks), control children's attitudes showed no significant changes. Treatment children showed significant attitude change on three scales, perceiving the teacher as liking them more, themselves viewing the teacher more positively and evaluating being good in class in a more negative manner. The authors suggest, in relation to this latter finding that the children possibly had a stereotyped view of themselves as 'bad', had learned to value deviant behaviour, and the programme had not made inroads on that view. They also point out that, contrary to many reseachers' expectations, attitude changes accompanying appropriate behavioural changes may not be wholly positive. This can also be seen as an important reason for trying to measure attitudes, although the findings may not be palatable. From baseline to the end of treatment (11 weeks), the earlier attitude changes were still present and had become stronger. In addition, their attitudes towards school in general and towards gym had become more positive.

These are intriguing findings, for a short period of intervention. Confidence in the results, however, has to be tempered by one important consideration. The controls in this study were not selected on the basis of .../

of deviant behaviour, but randomly, after the deviant children had been selected as targets. No data are supplied about pre-treatment strength of attitude, so important differences could have existed between the two groups which could to an unknown extent have contributed to the significant differences.

Dil & Gotts (1971) employed consumable rewards in an intervention focusing on arithmetic skills and arithmetic self-concept. The subjects were four 7-9 year-olds in one class, who were the lowest achievers, with the rest of the class functioning as controls. Rewards were available over an eight week treatment period, and, in addition, the curriculum and work groups were changed. Significant gains were reported for the experimental group in both arithmetic skills and arithmetic self-concept. However, there must be major reservations with this claim, since the groups were not equivalent at the pre-treatment stage, and effects cannot be clearly attributed to the reinforcement procedures since they are confounded with other changes.

(c) Studies of change in personality functioning

Changes in personality functioning have been explored even less than attitude variables in these interventions, despite the claims of theorists such as Eysenck (1967) for the relevance of personality factors for condition-ability. This neglect is presumably in large part due to the assumed linkage between personality measures and 'trait' theories which are now considered untenable (Berger, 1976). However, usage of such measures, as in the present study, need not imply acceptance of the 'trait' position. Responses on such measures can be taken as children's self-reports or statements about aspects of their own functioning without necessarily espousing the theory within which the measure has been embedded. Even self-reports, however, have been little used in behaviour modification with children.

As with attitude measures, no examples could be found in the behaviour modification literature of personality measures as dependent variables in interventions in regular secondary schools. One frequently cited study (Ward &..../

& Baker, 1968) did include personality measures in its assessment battery - the Draw-A-Person Test, in which drawings are taken to reflect aspects of the child's adjustment, maturity and self-image, and a projective questionnaire designed to measure attitudes toward school and feelings about self. The target children were four 6-year-olds, presenting a high frequency of disruptive behaviour, selected from three separate classrooms. Four control children were randomly selected from the same classes as the target subjects, and a further four controls were randomly selected from a fourth classroom. After a 6-week baseline, there was a seven week treatment phase, with teachers employing combined praise-and-ignore procedures. While favourable changes in target children's disruptive behaviour were reported, no significant changes on the personality measures were observed.

No examples could be found of studies where the specific measure employed in the present study - the neuroticism dimension on the Junior Eysenck Personality Inventory - has been used as a change measure in the context of a treatment study. With adults, Ingham (1966) found that change in clinical status, in a 3-year follow-up of psychiatric patients, was associated with a change in neuroticism, with the score of the improved group returning to the population mean. Hallam (1976) reported similar findings with phobics treated with behavioural procedures.

(d) Concluding note

It is evident that attitude and personality changes have received little in the way of systematic investigation in the context of classroom behaviour modification programmes reflecting, perhaps, the general narrowness of focus in measurement of changes. Additionally, the regular secondary school population appears to have been totally neglected. It has been suggested here that such measures could provide important information about possible correlates of changes in classroom behaviour. Their inclusion in the present study represents an attempt to fill this gap in the literature.

Having considered the impact of behavioural procedures in inducing change in a number of spheres of functioning we now turn to an examination of the extent to which such changes are generalised.

(v) Issues of Generalisation

(a) Introduction

The question of generalisation is a major one for all forms of behaviour modification. Three types of generalisation may be identified (Forehand & Atkeson, 1977): (1) Temporal Generality, or the maintenance of treatment effects after the termination of treatment; (2) Setting Generality, or the occurrence of treatment effects in settings other than the therapeutic one; (3) Behavioural Generality, or change in behaviours not the specific target of treatment interventions.

The present report was principally concerned with temporal and behavioural generality. The choice of treatment procedures was made with issues of maintenance in mind. Social reinforcement - teacher praise, approval, attention, etc. - was emphasised in preference to concrete reinforcement, on the assumption that it occurred more naturally in children's environments, and would be more conducive to maintenance of behavioural change. Provision was made for long-term follow-up to assess the persistence of changes. With regard to behavioural generality, the most direct measures of treatment procedures - the observational data - were supplemented by measures of functioning in other areas in order to monitor associated changes.

Although a number of studies demonstrate generalisation effects, these are in the minority (Kazdin & Bootzin, 1972; Marholin et al., 1976; Kazdin, 1977b; Stokes & Baer, 1977). Paradoxical as it may sound, generalisation <u>failures</u> are quite consistent with operant theory. If behaviour is regarded as a function of its consequences, then it will adjust to the contingencies prevailing in a given situation. Thus, if supportive reinforcement is withdrawn (as..../

(as with treatment termination or transfer to an unprogrammed setting) then the behaviour in question will probably decrease in strength. It is therefore important that generalisation issues be attended to in every intervention. A further irony is that while behaviour modifiers search for procedures that will enhance generalisation, the experimental design which has been central to their work - ABAB or reversal designs - may well have impeded them in their quest (Hartmann & Atkinson, 1973; Kazdin & Kopel, 1975). In this design, repeated treatment implementation is ideally associated with behaviour change, while withdrawal of treatment is associated with return to baseline levels. It is therefore difficult to demonstrate functional control and maintenance in the same study. Kazdin (1977b) suggests that 'the competing interest in clear experimental results and maintenance of behaviour change perhaps has led to sacrificing one for the other.' As will be evident from the discussion below, researchers have emphasised the demonstration of short-term reinforcement control, at the expense of longer-term maintenance.

(b) Temporal Generality

Despite the importance that is attached to the issue of generalisation and the plethora of methods advanced to encourage it (Wildman & Wildman, 1975) relatively few studies have directly addressed it. Studies of maintenance in classroom work are particularly limited in number. Some of the practical and theoretical problems which confront researchers contemplating follow-up, for example, have been examined in Chapter 3, and these have clearly deterred classroom workers as well as those in other fields. Kauffman et al. (1977) surveyed 152 separate experiments conducted in classroom or educational settings between 1968 and 1974, and found that 72% reported no follow-up data whatsoever. In the other 28%, follow-up was based on verbal report (28%), behavioural observations (54%) or standardised testing (18%). For studies relying upon verbal report, average follow-up duration was 107 days, but only 18 days for studies in which behavioural observations were employed. 36 of the 42 experiments in which follow-up was reported showed effective maintenance, but less encouragingly..../

encouragingly, a greater percentage of studies published prior to 1971 than those published since included follow-up data. The paucity of follow-up studies is not restricted to classroom studies in behaviour modification (Cochrane & Sobel, 1976; Keeley et al., 1976).

It is found in many classroom programmes that once reinforcing contingencies are withdrawn behavioural gains tend to dissipate (Birnbrauer et al., 1965; Kuypers et al., 1968) and behaviour change may not persist even into times of the day when programmes are not in effect (O'Leary et al., 1969; Meichenbaum et al., 1968; Wolf et al., 1968). In the O'Leary et al. study it was found that during a successful token programme run in the afternoons, the teacher delivered a high frequency of contingent social reinforcement for appropriate behaviour, whereas in the morning there was little social reinforcement delivered. The failure of afternoon behaviour to transfer to morning sessions appeared to be the result of different contingencies of reinforcement.

It will be recalled that several of the studies employing control group designs, reviewed in section (i) of this chapter, attempted to gather follow-up data. Safer et al (1981) conducted a 4-year follow-up and found that their treatment group had lower rates of withdrawal from school and suspension for disciplinary reasons. On achievement measures, the only significant differences favoured the control group. Wodarski & Filipczak (1982) experienced major practical problems with their 4-year follow-up and were restricted to self-report data. They obtained mixed results with no evidence of clear advantages for treated subjects. Rollins et al. (1974) report no follow-up data but indicate that despite the favourable impact of their programme, it had been abandoned in its entirety one year after the research team withdrew. The most encouraging report is that of Dickinson (1974) who found that pupils from a token reinforcement programme maintained an advantage on reading scores over controls two years after the programme ended. Specific criticisms of these studies have been considered earlier.

(c) Setting Generality

Few studies have been addressed to the issue of setting generality, despite the importance, in educational settings, of knowing whether programmes in one classroom have implications for children's behaviour in others, or whether improved behaviour in school transfers to the home. As with the issue of maintenance, generality of training effects across settings tends not to occur (e.g. Kuypers et al., 1968; O'Leary et al., 1969; Broden et al., 1970). Given the evidence as to the situation specificity of behaviour (Mischel, 1968; Wahler, 1969) these discouraging findings are not surprising. Where problem behaviour exists in two settings, intervention may be necessary in both. Thus, Wahler (1969) working with two boys showing deviant behaviour in home and school, found that modifying behaviour in the home did not result in corresponding changes in school. Change did not occur in the classroom until similar procedures of differential attention were applied there. Even with less dissimilar settings, transfer may not occur. Glavin et al. (1971) employed a token system to reduce disruptive behaviour and increase task-related behaviours in a 'resource' classroom for children of primary school age. The improvements obtained in this setting were not observed on return to the regular classroom, and behavioural and academic gains were not maintained at the 2-3 year follow-up (Glavin, 1974). O'Leary & Schneider (1977) similarly found that highly disruptive 6-year-olds who spent 8 months in a special class placement did no better upon return to their regular classes than similar children who attended regular classes for the entire year, with the exception of higher reading scores for the special class group. (It should be noted that maintenance and transfer issues are involved in many of these studies.)

(d) Behavioural Generality

The optimist who looks for temporal and setting generality also looks to a spread of effects to responses other than the ones targeted during a programme - whether this 'spread' occurs is the question of behavioural generality...../

generality. Behaviour modification has perhaps been excessively concerned with measuring only responses that have been the focus of treatment and the need for greater attention to multiple response evaluation has been expressed recently (Franks & Wilson, 1978). It is possible that because of this narrow focus in evaluation, many important changes have gone unmeasured and unrecorded. These changes may take the form of negative side effects as well as positive transfer to other behaviours. Prompted by the former concern, Ward & Baker (1968) found neither positive nor negative effects in WISC, projective questionnaire and Draw-A-Person scores following a programme involving changes in teacher attention with 6-year-old children.

Where positive 'spread' of effects is obtained, these are usually not programmed for. Kubany et al. (1971) for example, intervening to reduce disruptive behaviour in a 6-year-old boy , also found improvement in punctuality to class, although this was not included in the contingencies. Similarly, Twardosz & Sajwaj (1972) obtained gains in social interaction when reinforcing in-seat behaviour. The studies examined earlier which show achievement or IQ gains following reinforcement of attentional or task-related behaviours provide evidence of the wider effects that may ensue, but these results are not always obtained. While Ferritor et al, (1972) caution against hoping for by-products of reinforcement contingencies and recommend inclusion in the contingencies for any behaviour that it is desired to change, Nelson (1974) calls for a greater investigation of behaviours which may change collaterally so that interventions may become more economical. She suggests that behaviours which are compatible may collaterally improve, while those which are incompatible are likely to be inversely affected by reinforcement contingencies.

(e) Programming for temporal and setting generality

The message here seems to be clear, and it is conveyed in the frequently quoted but less frequently implemented suggestion that 'generalisation should be programmed rather than expected or lamented' (Baer, Wolf & Risley, 1968). As.../

As indicated earlier, generalisation was an important consideration in the design of the present study. Recently, methods for securing maintenance and transfer have multiplied as concern with these issues has grown. Some positive findings are emerging, and these can be considered in the light of the particular technique employed. The first of these involves developing behaviours which are likely to be maintained by the natural environment. Baer & Wolf (1972) suggested the benefits of a 'behavioural trap' - environmental support systems (e.g. praise and attention) which may result from altering particular child behaviours. If a behaviour can be developed even through artificial or contrived means, it may then be naturally supported and maintained by the consequences it elicits. Allen et al (1964) found that after increasing a withdrawn child's interactions, removal of attention did not result in relapse - the behaviour may have been maintained as a function of the 'trap' of social interaction.

Secondly, attempts can be made to substitute one programme for another. Walker & Buckley (1972), for example, investigated several ways of maintaining the effects of a special classroom token economy when children returned to their regular class. They compared 'peer programming' (subjects could earn points which were used to purchase reinforcers for the group), equating stimulus conditions (making conditions in the regular classroom as similar as possible to those of the special class), training the regular class teacher in behavioural methods, and a control involving simple return to the regular class without any special programming. During the 2-month follow-up period in the regular class both peer reprogramming and equating stimulus conditions were superior to the control. Cone (1973) has disputed these findings, suggesting that when baseline differences are taken into account, teacher training also shows positive effects. However, the maintenance of improvements is not in question. Walker et al. (1975), returning children to a regular class setting from a special education token economy classroom, compared a substitute programme with natural reinforcers (praise, grades) with a control condition with no programme in effect. Children with the substitute programme..../

programme maintained their behaviour better and when all programming was withdrawn, their superiority was evident at 4 months follow-up.

As a means of seeking generalisation across settings, Friedman et al. (1977) had their 13-14 year-old students, in a junior high school setting, carry special report cards into classes where their behavioural programme was not in operation. Ratings were recorded on these cards for the children's social and academic performance, and points were added to, or deducted from, their earnings in the base classrooms where the full programmes (token economy, interpersonal skills training, parent training) were being conducted. Relative to no-treatment controls, significant improvements were found in these 'non-programme' classrooms, and, interestingly, positive transfer was also found in a separate experiment in which the report card procedure was not employed.

The substitution of programmes, as a generalisation strategy, raises the question of how 'programming for generalisation' differs from what we ordinarily call 'treatment'. In some situations there may be no clear distinctions other than in factors such as intensity of therapeutic conditions, and degree of therapist or professional contact.

A third measure is to <u>fade the contingencies</u>. This strategy involves gradually removing or thinning out the reinforcement being provided. After controlling disruptive behaviour in an adjustment class with token reinforcement procedures, Drabman et al. (1973) gradually faded out the teacher rating by which point-earning was determined. During a final phase in which no ratings were made - unfortunately a brief 12 days - appropriate levels of behaviour were maintained. Similarly, Turkewitz et al. (1975) with eight disruptive 7-11 year-olds in an after-school tutorial setting, faded out back-up reinforcers and maintained behavioural improvement after they were completely withdrawn, albeit for the relatively short period of five days. Jones & Kazdin (1975) successfully faded out token procedures in a special classroom for retarded children and demonstrated that inattentive behaviour remained..../

remained at an acceptably low level three months after all explicit reinforcement contingencies had been withdrawn. Greenwood et al. (1977) obtained substantial maintenance over a nine-week period following termination of formal intervention procedures, by phasing out the specific behavioural programme materials, the classroom rules, and a class bar graph indicating daily progress. Results were significantly better than for control classrooms where no maintenance strategy was applied. Fading measures may to some extent involve phases of intermittent reinforcement, a technique that is traditionally recommended where maintenance is sought. It may also be implicit in the fourth strategy of delaying delivery of reinforcement. Greenwood et al. (1974) gradually increased the number of sessions before reinforcement could be earned and maintained appropriate behaviour at three-week follow-up.

A fifth method for seeking generality of treatment effects, and one which is gaining in popularity, is the use of self-reinforcement. Theoretically, if a child can become relatively independent of external reinforcement and can control his own behaviour, gains may not disappear when he moves out of the treatment setting or once programmes come to an end. However, it is often found that when attempts are made to replace external reinforcement by self-reinforcement, excessive leniency in self-reward may result (Santogrossi et al., 1973). This is not always the case, however. In the Drabman et al. (1973) study already cited, children had complete control over reinforcement administration in the successful 12-day phase when teacher management was eliminated. Bolstad & Johnston (1972) showed self-administered points to be superior to teacher-administered points in controlling disruptive behaviour, but the maintenance period was of only seven days' duration.

The importance of securing control group data when maintenance is assessed is shown in an important study with 8-10 year-olds by Kent & O'Leary (1976). A standardised 20-hour treatment programme involving the child, his parents and his teacher was evaluated, emphasising social reinforcement from the teacher..../

teacher and special home-based rewards and privileges contingent upon school performance. Relative to no-treatment controls, the treated children had improved significantly by the end of treatment, but at nine-month follow-up, gains by the control group wiped out these differences. However, treated children still showed better achievement scores and grades at follow-up.

Finally, the involvement of parents seems essential if it is intended that the effects of classroom-based programmes transfer into home. The extent to which parents control reinforcers for their children (e.g. access to toys, play activities, T.V., pocket money, etc.) underlines the potential impact of their involvement for generality of effects. Many studies have involved parents successfully to buttress school-based interventions (see Atkeson & Forehand, 1978 for review) though not necessarily directly addressing generalisation issues. These studies commonly involve conveying to the parents information concerning their child's behaviour in school and contingent rewards are then made available in the home. There is now a sizeable literature on parental involvement, reviewed by Johnson & Katz (1973); O'Dell (1974); and Graziano (1977) and its potential for prevention, as well as supporting and enhancing interventions conducted elsewhere, is enormous. As with other parent programmes, problems in enlistment, maintaining interest and modifying attitudes and behaviour remain (Chilman, 1973).

(f) Concluding note

To conclude this section, there is some minimal evidence for generalisation effects across time, settings and behaviours, although reviewers suggest that this is typically not the case, and that the balance of studies do not provide support for generalisation. A striking number of studies have not paid attention to generalisation issues, possibly influenced by expectations of narrow specificity of effects, as predicted by behavioural theory. The issues seem important enough, however, to merit planning for generalisation, and assessing such effects even in the absence of specific programming for them.

This completes the review of the effects of behavioural procedures in classroom settings. The next section deals with issues in the training of teachers to employ these methods.

(vi) Training teachers in behaviour modification

(a) Introduction

Given that a number of procedures exist which may be successfully employed by teachers to modify children's behaviour in the classroom, the question as to how teachers may most effectively be trained to apply these procedures needs to be addressed. The method adopted in the present study had three components: (1) a specially prepared document outlining the relevant behavioural theories and procedures, (2) three seminars, allowing more elaborate discussion of procedures, description of relevant studies from the literature and some limited modelling of techniques, (3) ongoing consultation, during which the details of individual interventions were worked out. This phase also allowed opportunities for supporting the teachers and encouraging them in their applications of the procedures. At a general level, the training of teachers in the use of behavioural procedures involves several issues: when teachers should be taught these procedures, how they should be taught, i.e. the best format for training, and what they should be taught, i.e. the most appropriate content. Each of these questions will be addressed in turn.

(b) When to train?

It is commonly agreed that the best time for teachers to learn about behavioural procedures is during their initial training, before inappropriate habits of responding are established (Thomas & Adams, 1971). Hawkins (1974) described one such approach at a college of education in Liverpool. This involved completion of a programmed study guide, practical training in observation and recording techniques using prepared videotapes, and some use of micro-teaching procedures.

However, the idea of including such exposure to behavioural approaches as a matter of course in initial training is far from a reality in this country, according to a survey conducted by Schwieso & Hastings (1981) of 114 institutions in England and Wales offering teacher training programmes, 54% responded to a request for information concerning coverage of behaviour modification in their courses.../

courses. Of the 252 courses being offered, 18 were reported to give behaviour modification 'incidental mentions', 36 gave specific lectures with or without film or video illustrations, and only 12 courses gave 'substantial' coverage with practical work, this being restricted to optional or specialist courses.

(c) How to train?

Such a picture suggests that, for many teachers, training in behaviour modification has to be conducted on an in-service basis. This leads to the second question of how such training is best conducted. This involves two issues: first, the general organisational format, and second, the specific methods chosen as the vehicle for training. Considering the organisational options available, training formats vary from individual work with teachers, through a variety of workshop presentations - which give the opportunity to work in groups- to more extensive graduate courses at college/university level. Many of the early, experimental and demonstration studies mentioned in the earlier review sections approached 'training' through informal discussion, advising on applications, and giving feedback on a one-to-one basis. Such a strategy has been effective in changing teachers' behaviour and has the advantage of allowing much individual contact and consultation and opportunities for personal encouragement, and support. However it is clearly uneconomical in terms of 'reach' - the numbers of children who are exposed to individual staff - and also in terms of professional time where successive individuals have to be trained. Some of the early reports, in fact, were quite discouraging for psychologists contemplating interventions in ordinary schools, as indicated in the opening Introduction. Abidin (1971) for example, suggested that psychologists would expect to spend approximately 30 hours assisting teachers with no previous experience in behavioural techniques to institutean individual programme and 150 hours to establish a token economy. A more realistic framework was proposed by Tomlinson (1972), conscious of time constraints, by which he worked through groups of teachers and reduced the psychologist's time required per case to, on average, 2.4 hours, and 4.2 hours for a token economy. At the other extreme, of some graduate level courses, for example,..../

example, the training programmes may extend beyond the constraints of ordinary school-sponsored in-service training by requiring more teacher time, effort, and money, and by taking place outside the confines of the school itself. Some variation on a 'workshop' format offers, perhaps, the best model for in-service training, where it can be practicably arranged within the regular teaching structure. By giving access to groups of teachers within a school, it offers the opportunity of creating a self-sufficient core of teachers who can consult with and support one another. Such a school-based model is currently being developed by Wheldall & Merrett (1982) in this country at the University of Birmingham. The training package is typically offered to six to eight volunteer teachers. Evaluative data are not yet available. Whichever training format is adopted, the question of specific vehicles for training efforts needs to be addressed. The following review of methods will also give examples of the various organisational frameworks referred to above.

While many training approaches have adopted multi-faceted procedures, a number of components may be identified which have been isolated in certain studies. Kazdin & Moyer (1976) identify five dimensions: (1) instructional methods, (2) feedback, (3) social reinforcement, (4) token reinforcement and (5) modelling.

Instructional methods, which include lectures, discussions and course work - generally didactic in format - are perhaps the most widely employed approach. Kazdin & Moyer suggest that the large numbers of instructional manuals for teachers which are now available reflects the popularity of the instructional approach. Despite the vogue for instructional workshops, there appears to be no reliable evidence that an instructional element leads to application of behavioural techniques rather than merely permitting acquisition of knowledge of such principles and methods.

A number of training programmes have been reported which rely heavily although some not exclusively, on an instructional format. Andrews (1970) working with eleven elementary school teachers, conducted four 90 minute weekly workshop sessions which considered reinforcement principles and their classroom application. Specific changes in teacher behaviour as a result of these sessions are not reported, but the teachers themselves described some reduction of problem behaviour within

their.../

their classrooms.

A training programme on a much larger scale has been developed by McKenzie and his colleagues (McKenzie et al., 1970) at the University of Vermont. Known as the Consulting Teacher Model, this approach takes the form of a 2 year programme leading to a Master of Education degree, andemphasises behavioural principles, measurement procedures and the systematic use of reinforcement techniques. While no systematic analysis of these training procedures has been conducted, Sherman & Bushell (1974) suggest that its general effectiveness is supported indirectly by the successful reports published by course participants, e.g. Hanley & Perelman (1971); Knight et al., (1971) and Lates et al. (1971).

A model similar in content to the Consulting Teacher Model is the Responsive Teacher programme developed by R. Vance Hall at the University of Kansas (Hall & Copeland, 1972; Hall et al., 1976). This is a graduate level programme which may be presented as an after-school university credit course or a district credit in-service training programme or workshop. Ten weekly 3 hour sessions is the usual format and, ideally, these are held on school premises after classes have ended for the day. Graduate students lead discussion groups of about 10 people and content covers recording and measurement procedures, experimental designs, examination of learning principles and research studies. Discussions are supplemented by lectures and films. Students are encouraged to apply what they have learnt by conducting experiments in their classroom which are subsequently discussed in the group sessions. It is in this way that Hall hopes that students will incorporate Responsive Teaching procedures into their actual teaching situation. There is a clear awareness of the need to translate knowledge into practice in this model and, indeed, to provide the combination of practical, theoretical and research skills which will enable teachers to continue to practice and build on what they have learnt after they have completed the course. The model's effectiveness is attested indirectly by the large number of studies which have flowed from it - Copeland & Hall (1976) report that some 60 Responsive Teaching studies.../

studies have been published, and some 2000 educators have been trained since the programme began. As with McKenzie's model, there is neither component analysis nor analysis of specific changes in teacher behaviour as a result of their training experiences. An in-service course, consisting of six 2 hour evening meetings, modelled after Hall's approach, was reported in this country by Harrop (1974). Although not fully documented, several successful experimental studies were claimed to have resulted from it. Harrop's course was for primary-school teachers. A similar course for approximately 100 secondary-school probationary teachers was conducted by McNamara & Harrop (1979). Teachers met on four occasions at fortnightly intervals, with each session being of two hours' duration. Sixty of the teachers who attended the final session completed a questionnaire, as well as writing up a case study on the effects of the intervention strategies. Only six teachers reported successful outcomes to the case studies. Success was adjudged on the basis of improvements in data collected by the teachers, and the teachers' positive comments. A further 30% of the teachers had a positive response to their intervention but no supportive data. Some 24% were either negative or neutral in their reports. McNamara & Harrop viewed this course as less successful than its primary-school counterpart and saw its brevity, difficulties in data collection, and a number of constraints within the secondary-school context as underlying its failure.

For interventions in ordinary schools, time-limited training models may need to be explored, given the constraints on teachers' time. Thus, the models of Harrop (1974) or Andrews (1970) may be more appropriate than the extended formats of McKenzie and Hall. As to content, Sherman & Bushell (1974) offer a rather jaundiced comment on this type of approach to training. They suggest that:

'For the most part the teaching procedures of these training systems appear to reflect the learning experiences of their designers. Hall, McKenzie and Andrews utilise relatively traditional academic practices that are generally assumed to correlate with the development of rather generalised skills' (p.443)

However, didactic procedures may be effective if supplemented with training in the situation in which the techniques are to be applied. Thus, McKeown et al. (1975) found that disruption in the classroom was reduced following teachers' participation in a laboratory group but not when only written information (a teacher's instruction manual) was provided. The laboratory group afforded opportunities for supervision and reinforcement of attempted applications; feedback as to the quality of performance; shaping of teachers' behaviour and modelling of performance described by the experimenters. The authors suggest that perhaps one reason for the small effects of the manual is that there was no real assurance that the teachers in the manual-only group adequately read the material. While this may well be a hazard in teacher training approaches which adopt this model, the omission by the experimenters in this study to ensure that the manual was read renders the validity of their comparison rather dubious, and weakens their conclusion.

Another experiment which demonstrates the insufficiency of an instructional approach alone is that of Bowles & Nelson (1976). In the first phase of this study, the performance of an experimental group of 13 teachers who were taught behaviour modification by means of an in-service workshop was compared with a no-treatment control group. The experimental group scored higher on a paper-and-pencil test of knowledge of behaviour modification, but there was no difference between the two groups on observed behaviour in class. In the second phase, half of the experimental group received two hours' training in their own classrooms during which audio cues were given to the teacher indicating when she should prompt, praise or make contingency statements to children. This procedure produced significant increases in this group in the trained behaviours relative to the teachers who did not receive such training. It is not clear from this study how long the effects were maintained, or the extent to which they were potentiated by the initial workshop exercise rather than being a function of the direct audio instruction alone. The authors suggest that the results are consistent with what is known of situation and response-mode specificity of behaviour, so that the test-taking behaviours reinforced in the workshop setting are not the teaching behaviours which are desired in the classroom. So in terms of specificity some workshop models may be geared to.../

to teach inappropriate behaviours in an inappropriate setting. Van Houten & Sullivan (1975) have also shown the effect of audio cueing in increasing teachers' praise rates, and also in maintaining these rates when the prompting was withdrawn.

The second element to consider in training approaches is the role of <u>feedback</u>. We will examine here some studies in which feedback is studied specifically, although the use of feedback is in fact implicit in some of the models already discussed, in the context, for example, of discussion of experimental applications of procedures and selective reinforcement of teachers' efforts. As Kazdin & Moyer (1976) suggest, feedback is inherent in virtually all forms of response consequation, but in the present discussion it will be taken to refer to the provision of information about the adequacy of performance or knowledge of results.

Feedback in training exercises usually takes the form of verbal or written reports of behaviour. Cooper et al. (1970), for example, used both these forms of feedback, verbal comments to one teacher and written notes to a second, referring to the teacher's frequency of praise and failure to praise appropriate child responses. Feedback was given every 10 minutes and successfully increased the percentage of time the teachers spent attending to the children's appropriate behaviour. It was found that the teacher who received written notes saved them until the end of the day, and, interestingly, the procedure was less effective with her, possibly with the delay in feedback. After feedback was eliminated, appropriate teacher behaviour began to decline, and data on maintenance were not collected. That the <u>immediacy</u> of feedback may be critical is also suggested by the failures reported in studies where graphical or verbal feedback is given at the end of the day (Cossairt et al., 1973; Rule, 1972; Saudargas, 1972) or every other day (Breyer & Allen, 1975).

Feedback may be given in other forms as well. Thomas (1972), for example, found self-recording by teachers of video-taped segments of their behaviour to be effective in inducing change. Saudargas (1972) replicated this and Rule(1972) has shown that self-recording of videotapes is somewhat more effective than daily instructions and feedback. Horton (1975) found that having teachers identify instances of behaviour-specific praise on videotape, combined with instructions to use praise, and audiotaped recordings..../

recordings of their classroom interaction as feedback helped to increase their rates of behaviour-specific praise. However, the effects were restricted to subject areas in which training was conducted. For example, training conducted in reading lessons did not generalise to maths, social studies and science. Rule (1972) also evaluated a rather aversive form of feedback - the experimenter praised the teacher for every 5 minutes of appropriate behaviour, but replaced her, taking over the teaching role, when her behaviour fell below the criterion level. This indirect intervention led to more appropriate teacher behaviour than either instructions and feedback or video recording, but its use would undoubtedly require good relationships and rapport with teachers. In the context of general teaching, Good & Brophy (1978) suggest that teachers are often unable to change some aspects of their behaviour in classrooms because they lack the conceptual labels which would generate awareness of what they are doing/not doing. Clear feedback on their behaviour is viewed as an important means of facilitating such labelling and providing the basis for analysing their actions and trying to change them.

The third dimension in teacher training to be examined is the use of social reinforcement. While classroom behaviour modification literature is replete with successful examples of social reinforcement techniques with children, there seems to have been a remarkable reluctance to consider the experimenter's or consultant's relationship with the teacher in terms of the principles of reinforcement theory. These principles can be viewed as operating at all levels in the consultant-teacher-child triadic model but relatively little cognisance has been taken of this.

Few studies are available in which conscious use has been made of social reinforcement to enhance the effects of training. Cossairt et al. (1973) found that merely telling teachers after each class how often the students attended and the amount of praise given produced no systematic changes in teacher behaviour. The experimenter than introduced comments of praise for the teachers' performance, and this markedly increased the teachers' use of praise in their classrooms. In addition,..../

addition, student attending behaviour increased under this condition. McDonald (1973) contrasted the effect of praising teachers for their selection of certain behaviour change strategies as opposed to simply telling them which strategies to employ. Teachers who received praise tended to show more 'supportive' behaviours and less 'desist' behaviours (reprimands or commands) than those who received consultation without praise.

These two studies demonstrate an awareness of the potential role of social reinforcement - employed in an explicit and systematic manner - at the consultant-teacher level. An interesting departure in the exploration of social reinforcement effects on teacher behaviour can be found in studies which examine another source of influence - the behaviour of children themselves. Conscious of the two-way, interactive nature of reinforcement, Graubard et al. (1971) moved from the traditional position regarding teachers as the sole source of influence in the classroom, and examined the implications for teacher behaviour of changing the children's behaviour. Modifying students' eye contact, and getting them to ask for extra help and make complimentary comments, led to dramatic increases in teacher praise and a decrease in negative teacher contacts. Sherman and Cormier (1974) reported essentially similar results after employing tangible rewards, unknown to the teacher, to change children's behaviour. The authors suggest, in fact, that from the training perspective, it may be more economical to try to establish effective teacher behaviour by focussing not directly on teachers themselves, but on the children.

These studies suggest that explicit use of social reinforcement, with the consultant or children as source, can be a potent influence on teacher behaviour. What is surprising is that so little use has been made of it.

As with social reinforcement, examples of the fourth dimension in training - the use of token reinforcement - are somewhat limited. McNamara (1971) compared the effect of token reinforcement with a response cost procedure (withdrawal of tokens) in altering.../

altering teacher attention to appropriate and inappropriate child behaviour. Points were exchangeable for cans of beer. The contingent delivery or removal of points was effective in altering teacher behaviour, especially so when a bonus contingency was in effect.

Another vehicle for teacher training efforts has been the use of <u>modelling</u> techniques. This procedure has not been evaluated extensively and only tentative conclusions may be drawn from the available studies, despite its potential impact in conveying explicit behavioural requirements, especially if modelling is conducted in the classroom where the teacher actually functions. It was in this setting that Ringer (1973) explored its use. An investigator (model) initially took major responsibility for the administration of verbal and token reinforcement while the teacher observed. The teacher was gradually 'faded in' and given progressively more responsibility for administering reinforcement, while the experimenter was 'faded out'. The results showed that the teacher was able to maintain substantial control after the model withdrew although her use of contingent attention was not altered. In an attempt to improve on the potentially limited effects of a didactic approach, Kubany et al. (1974) supplemented this phase by a procedure in which four teachers rotated roles in the classroom. One acted as teacher, one as an aide, one as an observer and one prepared the curriculum. Although adequate data are not presented on the effects of training outside this situation, the approach holds promise as a means of providing excellent opportunities for observational learning and forming discriminations about relevant aspects of teaching behaviour. An interesting, if controversial, variation of modelling was demonstrated by Brown et al. (1974). Working with one teacher who was somewhat resistant to a behavioural approach, a method of 'surreptitious' modelling was used. One experimenter went into the classroom to interact with the children and give them contingent praise and approval. Without giving...../

giving the teacher explicit instructions or even stating their intention, significant change in teacher praise and positive interaction with the children were observed. While the authors comment on the value of such an approach with teachers who do not openly favour behaviour modification, in that it avoids confrontation, some doubts may be raised as to the ethics of such an intervention, where the teacher is not aware of the goals of the exercise, and also as to the general applicability of such an approach, with its attendant problems of entry, and feasibility, in a regular classroom.

Kazdin & Moyer (1976) conclude that it may be necessary to employ a variety of procedures in teacher training rather than relying on a single approach. Modelling or role playing may ensure implementation more readily than didactic methods. They go on to suggest that 'more than any other single conclusion the.....evidence suggests that training teachers, or any other group, needs to incorporate a reinforcement system where actual behaviour in the situation is consequated'. (p. 192) Direct, on-the-job training may well be the most useful approach in changing behaviour, but it may not be a practical proposition in the regular classroom.

The difficulties that were noted earlier in generalising children's behaviour changes are, predictably, also present for the behaviour of teachers. The possible situation-specific effects of training have already been noted. Ensuring that training effects persist also poses problems. It would appear that provision of some form of reinforcement is as important in <u>maintaining</u> what teachers have learnt as in initiating it. Brown et al. (1969) and Cooper et al. (1970) have shown that withdrawal of reinforcement for teachers leads to their behaviour reverting to pre-training levels, with predictable consequences for children's behaviour. After training an entire school staff of 15 teachers, with impressive results with children's behaviour and school progress, Rollins & Thompson (1978) returned one year after withdrawing their support to find that the entire programme had been discontinued. They comment: 'Not only were the formal aspects of the program no longer in ..."

in use, but in-class observations indicated that the teachers had reverted to a more traditional style of teaching with concomitant increase in student disruptions and a reduction in student involvement.' (p. 236) These authors' attempted solution was to assist school principals to assume major responsibility for providing teacher training and implementing programmes. Their success in this, and that of Copeland et al. (1974) in a similar venture, suggests that the approach holds some promise in the maintenance of programmes since principals continue to be available to provide support and feedback and to advise teachers when problems arise.

(d) What do teachers need to learn?

Having examined the various ways in which training is conducted and the differing emphases adopted, the issue of what teachers should be taught needs to be considered more closely. The studies in the foregoing review indicate varying stress on the theory on which behaviour modification is presumed to be based, principles of application and actual practice. It seems reasonable to conclude from the evidence available that knowledge of theory or principles does not ensure practice, and that some form of more direct training, support, or consultation in situ may be necessary to encourage implementation.

One possible off-shoot of this realisation has been an over-emphasis on practice and application of techniques at the expense of conveying knowledge and understanding of what the behavioural approach entails more generally. Berger (1979) has cautioned against what he calls a 'mindless technology, the use of techniques, divorced from their theoretical base and applied framework' (p. 418). These risks, he suggests, arise from a combination of the apparent simplicity of the techniques, the enthusiasm to pass them on to non-psychologists, particularly teachers, and an inadequate approach to consultation where techniques are applied in a simplistic fashion, and provision is not made for adequate monitoring and follow-up. In a related vein, Stein (1975) criticised brief workshops with no follow-up because they..../

they may not result in proper application of the principles taught. Moss & Childs (1981), considering workshops in which each teacher participant had to modify one particular problem in one particular child, question whether the teacher is then equipped to deal with other problems with other children in other conditions. Given the likely situation-specificity of such learning, one would suspect not.

It seems necessary to go beyond narrow presentation of techniques which can degenerate into a parade of 'tips for teachers' to provide a conceptual grounding in behavioural principles of problem development and change and to convey a grasp of the wider relevance of what is being learnt. The ideal, of course, is a balance between theoretical understanding and practical experience. It is possible for treatment of theory to be overdone as well, especially where it is of dubious relevance as Schwieso & Hastings (1981), for example, argue in the case of behaviourism as an underpinning for behavior modification. It is also perhaps best to avoid the use of jargon and any insistence on the use of precise 'scientific' language in the training exercise (Tomlinson, 1972). The terminology of behaviour modification may be off-putting and alienating to some. Woolfolk et al. (1977) report an interesting experiment demonstrating the impact of the labels used to describe behaviour modification. Teachers who were shown identical videotapes of a teacher using reinforcement methods evaluated the teacher and the teaching method more favourably when the videotape was described as illustrative of 'humanistic education' than when it was labelled 'behaviour modification'. The terms in which the approach is couched can therefore mnake an important difference to its evaluation and acceptance.

(e) Concluding Note

In conclusion, training within the regular school framework needs to entail some variant of an in-service approach which is compatible with other ongoing teaching commitments, and which does not make unrealistic demands on time, organisation and resources. The time factor is a crucial one - both for the teacher and for the psychologist - and at the time of planning the..../

the present project, most of the available models seemed either excessively time consuming or organisationally too demanding (e.g. degree courses, extended in-service training). A more limited, but effective model had to be sought. Although research since then has pointed to the inadequacy of simply conveying knowledge about behaviour modification and indicated the need for a practical component in training, no universally accepted training model has emerged. The experimental evidence for most is partial, and open to qualification. In the present training approach, the first two components (document and seminars) sought to convey the necessary knowledge, with a balance between theoretical understanding and practical knowledge of procedures. The third component - consultation - was intended to supply the bridge between knowledge and application.

(vii) <u>The Working Hypotheses</u>

Having concluded the survey of the classroom intervention literature for behaviour modification, it would seem useful at this stage to outline the working hypotheses of the study. These form the basis of the research hypotheses which will be formally stated and operationalized in Chapter 5, in Part 2, which is concerned with the description of experimental methods and treatment procedures.

The principal aims of the main project were to identify maladjusted school children and to apply and evaluate different ways of conveying help (Kolvin et al., 1981). The ultimate concern was to improve the functioning of these children in a number of different areas - emotional, behavioural, social and academic. The hypotheses of the main project covered a range of issues which are not the concern of the present report, in which the hypotheses are elaborated from the perspective of behaviour modification.

The main working hypothesis here was that by training teachers in the use of behaviour modification procedures, significantly better improvement in a number of areas of functioning would occur in the behaviour modification regime than in treatment comparison regimes or in untreated controls. As reflected in the foregoing review of intervention studies, 'improvements' could be sought in a number of areas through the medium of behavioural procedures. However, the apparently limited application of the principal procedure to be adopted in this study (social reinforcement.../

reinforcement) with the secondary-school population was a consideration that encouraged caution. In general, the position was adopted that the exploration of social reinforcement was well justified. Its neglect appeared to be based on an assumption of non-effectiveness rather than on any established pattern of negative findings.

The behavioural regime was concerned to effect changes in the following areas: (a) classroom behaviour, where it was hypothesized that behavioural procedures could have a positive impact with the range of child behaviours in the classroom which could be regarded as hindering the effective functioning of either individuals or the larger group of learners. This might range from the intrusive forms of misbehaviour, such as aggression, disruptive noise, or talking, to problems such as inattention, distractibility or withdrawal and inhibition in classroom activities, which, while perhaps impinging less on other classroom members, may nevertheless significantly impair learning for the individuals concerned. Improvements in the sphere of classroom behaviour could be reflected in a lessening of acting-out behaviours and in an increase in task-related activities.

(b) social interaction. The second working hypothesis was that teachers could use behavioural procedures to improve the functioning of those children who were experiencing interpersonal difficulties within the classroom group. The relationship between social competency and life adjustment in a number of spheres appears to be well-established. Although this association is mostly based on correlational studies, there seems little doubt that improvements in interpersonal functioning can have important implications for children's current and future adjustment. Involving teachers in such interventions would seem to have strong face validity.

(c) academic functioning. It was hypothesized here that the use of behavioural procedures to increase task-relevant behaviours would be associated with gains in academic functioning. The assumption was made that improvements in attention and application to assigned tasks, listening to instructions, concentration, and so on, would have significant implications for academic performance.

In addition to these basic working hypotheses, it was also hypothesized that changes...../

changes would be effected which would reflect (d) generalization processes. This involves, first, behavioural or response generalization and, second, generalization in time, or response maintenance. Where the former generalization issues are concerned, it was noted in the literature review that little attention has been paid in classroom behavioural research to changes in behaviour or aspects of functioning other than the direct targets of intervention. It was argued that a possible 'spread' of effects beyond immediate targets could be explored usefully, and the utility of employing personality and attitude measures as dependent variables was suggested.

Where the second aspect of generalization - maintenance in time - was concerned, it was hypothesized that by attending to maintenance issues at the time of planning and applying treatment procedures, obtained changes could be maintained beyond the treatment phase..

Hypotheses were not advanced in the present exercise to tap a third aspect of generalization - setting generalization. Some consideration was given to this in the main study where some home-based measures were included (Kolvin et al., 1981).

The manner in which these working hypotheses were operationalized and formally stated, and the experimental design which was employed to test them, will be fully described in Chapter 5.

Before these details are presented, the next two sections will examine briefly the intervention literature pertaining to the two treatment comparison regimes.

B. Consultation and Casework Studies

Introduction

This section of the review will deal with the literature relevant to the parent counselling/teacher consultation regime. Given the status of this condition as a comparison method for the main treatment intervention, rather than as a method of intervention to be fully explored in its own right, the review will be brief.

The review will reflect the two basic components of the approach: first, consultation with teachers by social workers aimed at expanding and enhancing the pastoral role of the teacher, and second, a short-term social work approach with parents..../

parents to promote both parental awareness and appreciation of the way family factors may influence children's behaviour. Attempts to link home and school complemented these aims.

(i) School based consultation and case work

Consultation may be defined as collaborative problem-solving between a specialist (consultant) and one or more persons (consultees) who are responsible for providing some form of assistance to another (client) (Medway, 1979). Consultation models with which school social workers have identified are those influenced by the theories underlying Mental Health Consultation developed by Caplan (1970) for work with teachers. This model of consultation stresses the interpersonal qualities and skills of the consultant such as attentive listening and the ability to empathise.

While there is an abudance of studies of consultation that is school-based, studies of social workers' involvement in such work are very limited in number. There have been two major reviews of the effectiveness of consultation in recent years, by Mannino & Shore (1975) and Medway (1979). The latter was specifically concerned with school-based work. These reviews are concerned with three major areas of outcome. First, changes in the knowledge, skill, attitude or behaviour of the consultee; second, change in the client group (the children) with whom the consultee works, and third, change in the organisational structure or system. Of the studies between 1958 and 1972 reviewed by Mannino & Shore, 20 were school-based. Most were concerned with change in more than one area. Of the 14 which sought change at the level of the consultee, 9 showed positive results. Eleven studies were concerned with change at client level and six of these reported positive findings. Of the 3 which sought organisational or system changes, only one reported success. Although all but one of these studies used a control or comparison group, the authors point out that a common methodological failure was that of not separating consultation effects from the effects of mere attention..../

attention. It is also evident that greater success was achieved in effecting change at consultee level than at the level of the client, which is a much more demanding test of such an indirect interventional model. One of the studies reviewed, Schmuck (1968), though demonstrating significant cognitive and attitudinal changes in consultees, also found that there were no corresponding behavioural changes, and no changes at the level of the children. Such findings raise important questions concerning the implications of change of certain kinds at consultee level for client change. The question of which level of change should be given primacy is unclear. Cowen (1978) however, has forcibly stated that even if 'consultees ...are shown to have enjoyed a programme, learned a lot from it, and to have developed more favourable mental health attitudes, it cannot be assumed that those changes lead to more effective helping, or growth-supporting practices.' (p.797) These comments clearly reflect a desire for more stringent evaluative criteria.

The survey by Medway (1979) dealt with studies published between 1972 and 1975. Of the 29 studies reviewed, eight investigations reported consistent positive effects resulting from consultation, that is, positive results were found on all the measures employed. Six of these studies involved behavioural consultation. An additional 14 studies which utilised multiple dependent measures obtained desired effects on some indices but not on others. Overall, therefore, consultation was at least partially effective in 76% of the studies reviewed. Taking those forms of consultation which may be identified most closely with the model employed in the present study - mental health consultation and organisation development (process consultation) - yields ten studies. Only four of these showed at least partial effectiveness and only one reported a positive outcome at client level. Medway points to a number of design problems which limit confidence in the findings of individual studies. The most serious is that eleven studies, of which ten reported positive effects, failed to include a control or comparison group. Other concerns are lack of control of bias in assessments, and failure to deal with attention or placebo effects. An interesting feature of the studies surveyed is that most of them assessed change at consultee level (79%), less at client level (41%) and only 7%.....

7% at system level. Only two studies examined change in client achievement and neither reported positive findings. These two surveys, therefore, offer only much qualified support for non-behavioural school-based consultation. It would appear that evidence for effectiveness at client level is very limited, and this is the level of change that is of most relevance to the present study.

One of the deficits in this area of research addressed by Medway was the lack of comparative studies of various consultation approaches. Since his review, two studies have appeared (Jason & Ferone, 1978; Jason et al., 1979) comparing behavioural and process models, and behavioural, process and ecological models respectively. As described in an earlier section, the behavioural model tended to be the most effective and the only positive finding for process consultation was an increase in achievement scores in the latter study. Both studies, however, were marred by methodological limitations.

While the above reports are not particularly encouraging, somewhat more positive findings of relevance to school-based social work have come from work conducted in the U.K. by the Central Lancashire Family and Community Project (Rose & Marshall, 1974; This project ran from 1965 to 1973 and involved placing teacher/social workers within four secondary schools. Each worker was based in a single school as a staff member, and had a small teaching load in addition to the primary counselling or social work function. They were also expected to provide a home-school link and to engage in family social work where necessary. The study is somewhat difficult to evaluate since the nature of the help provided is not clearly specified. It is defined by Marshall & Rose (1975) as 'any kind of work which operated through personal relationships to help children and their families to solve or adjust to personal and social problems'. (p.3) It is hard to imagine what kinds of helping procedures would <u>not</u> fall under such an umbrella. Furthermore, in the words of Marshall and Rose, 'workers and schools were allowed complete freedom to pursue their own ends' (p.3) so that case selection was left to the social worker and school staff rather than organised by a uniform procedure. This created difficulties in securing controls within the treatment schools, since children not selected as cases would not be as poorly adjusted. Controls were sought in the other schools not involved in treatment, but.../

but matching problems arose here also, in that while the initial delinquency rate of the control schools was similar to that of the combined experimental schools their initial maladjustment rates were different. Individual matching of subjects was therefore resorted to, so the possible biases attendant on non-random assignment could not be avoided. Over 500 children were involved in the programme over the period of operation. Just over four hours, on average, was spent on each case, with over half this time being devoted to individual interviews with the child. Family consultation took up about 22% of the time, teacher consultation about 13% on average. The distribution of involvement is therefore quite different from that of the present social work regime which did not include direct contact with children. Treated children improved on measures of general behaviour, school attendance and social adjustment, but did not differ from controls in terms of changes in personal adjustment scores. In addition, delinquency rates in boys in experimental schools reduced substantially relative to controls. Having begun with similar rates in the first year, by the fourth year there were only three-quarters as many delinquents as in the control schools, and by the age of 17/18 (amongst former pupils) only two-thirds as many. These outcomes need to be seen in the light of the design problems cited earlier. This study at least suggests that school social work can be effective and perhaps the permanent appointment of workers to the schools was an important contributory factor. A strong link or attachment to schools, as in the present study, may therefore prove of considerable value.

(ii) Casework studies

Fischer (1976) has reviewed studies of social casework up to 1975 which included either an untreated control group or a comparison group treated by nonprofessionals. Social casework was defined as 'the services provided by professional caseworkers' (p.11). This definition was adopted in order to provide the broadest possible base for evaluating casework, by including whatever services professional caseworkers deemed necessary. The assumption is made, therefore, that casework is not a unitary or homogeneous phenomenon but covers a diversity of methods. Fischer identified seventeen studies which met his criteria for selection. Of these, ten involved children.../

children or adolescents and their families. Two studies (Lehrman et al., 1949; Levitt, 1959) dealt with general child guidance clinic cases covering a range of presenting problems. One study was concerned with 'socially disadvantaged, intellectually superior children whose problems were viewed as a function of a pathological environment' (McCabe, 1967). The remaining seven studies were concerned with delinquent or anti-social behaviour problems (Powers & Witmer, 1951; Tait & Hodges, 1962; Miller, 1962; Craig & Furst, 1965; Meyer et al., 1965; Berleman & Steinburn, 1967; Berleman et al., 1972). The services provided in these studies included a range of individual, group and environmental activities designed to modify the social functioning, feelings and attitudes of the clients either directly, or by affecting those in their immediate living environment.

Fischer's conclusions are most discouraging. All but one of these studies showed no differences between the experimental and control groups. The one exception (Lehrman et al, 1949) claimed superior gains in the treated group when children were rated as 'improved'. However, Fischer points out that when those rated 'potentially improved' are also considered, the differences between the groups disappear. He also criticises the use of defector controls in this study. For Fischer, therefore, none of the studies examined provided clear evidence that any form of social casework was effective.

These conslusions have not gone unchallenged. Thus, it has been suggested that the failure to provide more specific and focussed definitions of treatment processes renders an analysis of the effectiveness of 'casework', as a vague and global concept, quite meaningless (Hudson, 1976). Fischer is also accused of criticising studies which claim positive effects but not those showing negative effects (Cohen, 1976) with the inference that if studies are inadequate, they cannot be used to support or attack casework.

Despite these, and other criticisms, the absence of sound supportive evidence is a matter for concern. Fischer's review was restricted to American studies. The balance is unfortunately not redressed in the U.K. Thus, Berg et al. (1978) showed that repeated adjournment procedures for persistent truants (involving frequent attendances at.../

at court by the child and his parents) proved more effective than supervision orders under which the child would be seen at intervals by a social worker. Not only were truancy rates significantly reduced for the adjournment group, but they also showed a lower incidence of crime.

A number of criticisms of social casework research are similar to those levelled at psychotherapy research, e.g. diffuseness of goals, the masking of individual variations in response to treatment, and inadequate tailoring of therapy process to particular problems. There is ample scope for more focused research in the casework field. One such study is that of Reid & Shyne (1969) in which different varieties of casework were contrasted. They studied 120 families presenting difficulties in marital or parent-child relationships and compared short-term casework (up to 8 interviews within a 3 month period) with an open ended approach involving up to 100 interviews over an 18 month span. Systematic measures of various aspects of personal and family functioning showed a significantly higher proportion of favourable outcomes for the short-term casework group. Examination of therapy process suggested that the greater benefits of the brief approach lay in its better definition of goals and its clearer treatment focus.

(iii) Conclusion

Reviews of the effectiveness of school-based consultation and social casework point to little supportive evidence for either form of helping intervention. Evidence for change at client, as opposed to consultee, level, is particularly lacking. Methodological problems create difficulties in interpreting studies with both positive and negative findings. School-based interventions by social workers have been little researched but the limited findings available suggest that such an involvement merits further study. Where casework is concerned, the short-term approach, with clearly defined objectives and treatment focus, offers perhaps the most promise.

C. Group Counselling Approaches with Children

Introduction

This..../

This section will present a brief examination of studies of group counselling with children experiencing problems within the school context. As with parent counselling/teacher consultation, this approach was employed as a comparison treatment. In order to set this review in context, some consideration will be given to the nature of groups and the assumptions underlying how group approaches are perceived as helpful.

(i) The nature of groups

It is generally accepted that a human group is more than just a collection of individuals. Freud (1922) suggested that a leader was of particular importance in a group, viewing him as a father-figure for group members. The sharing of a leader leads to recognition by members of their common identification and dependency which generates a feeling of group identity. Sherif & Sherif (1969) define a group as a 'social unit consisting of a number of individuals who stand in role and status relationships to one another, stabilized in some degree at the time, and who possess a set of values or norms of their own regulating their behaviour, at least in matters of consequence to the group' (p. 131). This definition conveys the notion of a sense of group identity which underlies feelings of 'belongingness'. This implies that the same people are involved with one another over a period of time.

(ii) How groups can be helpful

Group approaches can be viewed as sharing the same objectives as individually-based methods, in seeking to help clients with the range of presenting adjustment problems. However, there are a number of differences between these approaches. Hopson (1982) suggests that the group situation can provide immediate opportunities to try out ways of relating to individuals, and is an excellent way of providing the experience of intimacy with others. The physical proximity of members to one another can be emotionally satisfying and supportive. In addition, members of a group not only receive help themselves but also help others. In group settings, members often discover that.../

that other people have similar problems, and that they are not alone in their suffering. A number of additional potentially therapeutic features of groups have been outlined by Yalom (1975). These include (1) opportunities for children to learn social skills such as turn-taking, delaying gratification, persuasion, etc. (2) group cohesiveness, where the mutual attraction of members for each other is seen as making them more open to influence from others,more accepting of others' hostility and placing greater value on group goals than are members of less cohesive groups, (3) a number of existential factors, including issues such as the need to recognise that life is at times unfair and unjust, that there is ultimately no escape from some of life's pain, and that people must take responsibility for the way they live their own lives, regardless of how much support and guidance they get from others. These issues are especially relevant ones for young adolescents at a stage in development when responsibility for one's own life is becoming an increasingly salient concern.

The group intervention in the present study employed a non-directive counselling approach, marrying the helping potential of groups with the creation of therapeutic conditions by the therapist. The non-directive stance is usually associated with Carl Rogers (1952), who emphasised the facilitative conditions of empathy, positive regard and genuineness as necessary and sufficient conditions for change. Much research has been conducted within this framework examining the relationships between these conditions and the outcome of therapy (Truax & Mitchell, 1971).

<u>(iii) Group counselling with children's school-based problems</u>

A substantial number of controlled evaluations of group counselling have been conducted with the range of school-based problems with which the present report is concerned. The bulk of these research studies has been addressed to problems in the following four areas: (1) academic under-achievement, (2) social isolation or withdrawal, (3) poor self-concept and (4) problem behaviour, particularly behaviour considered disruptive in the classroom.

The samples in these studies tend to be drawn most often from teacher nominations of problem children with little in the way of further psychological evaluation which would help to establish the representativeness of the children selected. Other common selection methods are scores on sociometric measures and on achievement tests.

In a review of such studies (which also included studies carried out with psychiatric as well as school populations) Abramowitz (1976) suggests that the available evidence regarding group counselling outcome is inconclusive, and somewhat discouraging. About one-third of the studies she reviewed showed positive results, one-third produced negative findings and one-third produced mixed results. A disquieting feature noted by Abramowitz is that a disproportionate number of the negative verdicts are contained in unpublished theses located in Dissertation Abstracts International. She speculates that since negative findings are often difficult to publish, there may be many other negative non-dissertation outcome studies which have gone unreported. (This possible publication bias is not of course specific to group counselling studies.) It is also pointed out that the proportion of positive to mixed to negative findings remains fairly constant regardless of the quality of the investigation and the domains of measurement which are tapped.

Examining studies employing control or comparison groups with school-age children in the four problem areas identified earlier, the balance between positive and negative outcomes is remarkably consistent. In the group of studies dealing with academic achievement, improved reading performance on formal tests was obtained by Fisher (1953), Shouksmith & Taylor (1964), Deskin (1968) and McCollum & Anderson (1974), while academic gains were reported in teacher ratings by Mann (1968) and in grade point averages by Randolph & Hardage (1973). These latter measures are composite continuous assessment scores that are available routinely. The reliability of such measures tends to be ignored, as is the question of whether the assessment might be contaminated by knowledge of the child's participation in a special project..../

project. On the negative side, Strickler (1965), Cheatham (1968) and Moulin (1973) found no improvements in formal achievement measures, as did Clements (1963), and Winkler et al. (1965) in grade point averages. Many of these studies gave very limited information about therapy processes, so that it is not always clear with what particular outcomes might be associated. Describing therapy processes is admittedly difficult, and actual practices may differ from what a particular theoretical framework might predict. Thus, Lieberman et al. (1973) found little relationship between the professed type of therapy, in encounter groups, and the results of objective observation. Given the shortcomings of description in the above studies, the more successful studies did appear to have a sharper focus on the solution of academic difficulties.

Another group of studies focussed on peer relationships, with children selected on the basis of sociometry scores as having few friends. Improvement on sociometric scores has been demonstrated in three studies, by Kranzler et al. (1966), Bevins (1970) and Thombs & Muro (1973). However, a further three studies showed no therapeutic effect (Biasco, 1966; House, 1971; McBrien & Nelson, 1972). Work geared to improving self-concept has been similarly inconclusive with positive findings by Crow (1971) and House (1971) balanced by the mixed results obtained by Mann (1968) and the negative outcomes of Clements (1963) and Hugo (1970). The same pattern emerges with disruptive/ inattentive behaviour in the classroom. Positive outcomes were recorded by Barcai & Robinson (1969) on teacher judgements, in behavioural observations by Hume (1967) and on teacher behavioural ratings by Crow (1971). Barcai et al. (1973) found a mixed outcome on teacher ratings, with responses inconsistent across teachers, while Seeman et al. (1964) and Hubbert (1978) found no improvements on teacher ratings.

To conclude this brief survey, a few general comments on the studies can be offered. First, many of the reports are based on small numbers which render the attempts at group comparison of doubtful value. Second, none of.../

of the studies made adequate allowance for initial pre-treatment differences between groups. Third, follow-up, beyond end-of-treatment measures, has been almost totally neglected.

(iv) Conclusion

It is evident that there are no consistently clear trends in the group counselling literature that has been surveyed here. However, although negative findings abound, it does seem that group counselling can help children, and that positive outcomes have been recorded in all the problem domains considered here. While recognising these encouraging possibilities, it must also be recognised that the methodological limitations which are in evidence in other areas of therapy research are equally prevalent here.

This concludes Part 1 of this thesis, in which the various strands in the theoretical background to therapeutic interventions with maladjusted children have been explored, with particular emphasis on behaviour modification as a treatment approach. Part 2 of the thesis will describe the methods employed in the writer's own investigation.

PART 2

METHOD

CHAPTER 5

DESCRIPTION OF EXPERIMENTAL PROCEDURES

DESCRIPTION OF EXPERIMENTAL PROCEDURES

This chapter will present details of the experimental procedures employed in the study. It will describe the following: the experimental design of the study; the research hypotheses; screening procedures; allocation of children to regimes and how these regimes were organised; additional measures employed to supplement the information from screen instruments, and the timing of assessments. The next chapter will be devoted to describing the three treatment regimes.

(i) Overview

The work described in the present report formed part of a larger scale research project funded by the Department of Education and Science, reported in Kolvin et al (1981). The aim of this study was to examine the effectiveness of different forms of help provided for maladjusted children identified within schools. The present report is specifically concerned with an examination of the effectiveness of behaviour modification as one of the treatment modalities. The children who participated as subjects were selected on the basis of screening procedures conducted in the first year of six junior high or secondary schools, and were then randomly allocated to treatment or control groups. Evaluation of the effectiveness of treatment was carried out by means of a variety of measures gathered from teachers, peers and the children themselves. These evaluations were conducted before and after treatment and at two subsequent points, spanning a follow-up period of three years from first contact. In the following sections, a detailed account will be given of the methodology and organisation of the experimental study.

(ii) Experimental design

The overall design consisted of four regimes - three treatment regimes and one no-treatment regime (maladjusted controls). Each regime was represented in six separate schools and the programme was implemented across two years. Table 5(1) outlines the basic design.

Table 5(1). Basic design of research project

Regime	Maladjusted Controls (MC)	Parent Counselling/ Teacher Consultation (PC)	Behaviour Modification (BM)	Group Counselling (GC)
Number of cases	92	83	74	73
Project Year	1	1	2	2

The present report focuses on an examination of outcome comparing behaviour modification with untreated maladjusted controls, parent counselling/teacher consultation and group counselling. Briefly, in the behaviour modification regime, the writer functioned as a consultant to all the teachers involved in this regime across the six schools. These teachers applied behavioural procedures on a class-wide basis but with a special emphasis on the selected children in their class. In the parent-counselling/ teacher consultation regime, six social workers, one attached to each school, functioned similarly as consultants to teachers, but with a focus more on psychodynamic aspects of children's behaviour than on their overt behaviour as in the behaviour modification regime. In addition, this regime had a further component involving home-based casework with parents. In contrast to these two regimes, in which the children were 'treated' in their regular classrooms, children in the group counselling regime were withdrawn from their classes for discussion groups, which were run on non-directive Rogerian lines by the same six social workers who had conducted the parent-counselling/teacher-consultation regime. The three regimes will be fully described in the next chapter.

Why were these three treatments selected? At a general level, all three were school based - the parent counselling/teacher consultation regime partially so. They also..../

also had the potential of reaching large numbers of children. These were two important concerns in the basic thrust of the project as a whole. At a more specific, individual level, the rationale for selecting behaviour modification has been presented fully in the Introduction. The two other regimes both provided the opportunity of deploying social workers in schools in innovative, but very different, ways. Both these regimes also involved a close concern with children's feelings and internal events, in contrast with the focus on overt behaviour in the behaviour modification regime.

Parent counselling/teacher consultation was particularly attractive as a treatment comparison regime in having a teacher consultation component - as with the behaviour modification regime. This, it was hoped, would enable some examination of the extent to which consultation, per se, could account for the outcome observed. Consultation, however, cannot be regarded as "controlled" in any strict sense for a number of reasons. First, no attempt was made to equalize consultation time. Second, there were six consultants engaged in the PC regime, one attached to each school. In the BM regime, one consultant operated across all six schools. Third, more teachers were involved in consultation in PC than in BM. Finally, PC had an additional component of home-based work which BM did not, so the comparison is not simply directed at two differing models of teacher consultation. Group counselling offered an equally attractive contrast condition, not only in offering a very different orientation (non-directive, emphasising Rogerian principles) to the behavioural regime, but also in representing a <u>direct</u> form of treatment, with the therapist in fact-to-face contact with the children. The other two regimes can be regarded as <u>indirect</u> models in operating through consultation. So the three treatments offered several dimensions of comparison and contrast: overt behaviour versus psychodynamic or 'feeling' focus; indirect versus direct mode of contact; teacher-based versus non-teacher based.

In addition to the control groups in the schools in which treatments were being conducted (the within-school controls), controls were also sought in two further schools in which no treatment was given (between-school controls). The aim../

aim of seeking between-school controls was to control for possible 'contamination' in the case of untreated controls located in the same schools as treated children. However, it was established that differences between the two 'between-schools' and the others on measures of characteristics of the school environment were such that it would have been inappropriate to rely on between-school data. These were not therefore included in the present analysis. It should be noted that while 'contamination' of the within-school controls was possible, it would tend to reduce differences between treated and control children, rather than magnify them. It would therefore tend to militate <u>against</u> establishing a positive effect of treatment.

(iii) <u>The research hypotheses</u>

As stated in Chapter 4, the main working hypothesis was that by training teachers in the use of behaviour modification procedures, significantly better improvement, in a number of areas of functioning, would occur in the behaviour modification (BM) regime than in treatment-comparison regimes or in untreated controls.

As indicated in the description of the experimental design, these comparison and control regimes are: parent-counselling/teacher-consultation (PC) and group counselling (GC) as the treatment-comparison regimes, and untreated maladjusted controls (MC) as the control regime.

The formal statement of the research hypotheses follows. The measures referred to here, which were employed to test the hypotheses, are described in sections (iv) and (vii) below.

(1) There will be significantly better improvement for BM relative to MC, PC and GC on teachers' ratings of behaviour in school and classroom at (a) end of treatment, (b) midline assessment and(c) final follow-up. The relevant measures here are the Rutter B2 Teachers' Rating Scale (at points (b) and (c) only) and the Devereux Elementary School Behaviour Rating Scale.

(2) There will be significantly better improvement in the BM regime than in MC, PC and GC on sociometric measures at (a) end of treatment, (b) midline and (c) final follow-up assessments.

(3) There will be significantly better improvement in the BM regime than in MC, PC and GC on personality self-ratings at (a) end of treatment, (b) midline and (c) final follow-up..../

follow-up assessments. The measure analyzed here was the neuroticism dimension on the Junior Eysenck Personality Inventory.

(4) There will be significantly better improvement in the BM regime than in MC, PC and GC in measures of ability. The relevant measures here were verbal and non-verbal ability scores on the General Ability Test (NFER).

(5) There will be significantly better improvement in the BM regime than in MC, PC and GC in reading comprehension at (a) end of treatment, (b) midline and (c) final follow-up assessments. The measure employed here was the Reading Comprehension Test (NFER).

(6) There will be significantly better improvements in the BM regime than in MC, PC and GC in school-related attitudes at (a) end of treatment, (b) midline and (c) final follow-up assessments. The relevant measures here were the ten attitude scales within the Barker-Lunn Attitude to School Questionnaire S - 7.

(7) Observational ratings of classroom behaviour in the BM regime will show an increase in task-related activities during the period of intervention.

(8) Increases in task-relevant behaviour within the BM regime will be associated with gains in reading comprehension as assessed at end of treatment.

The statistical methods employed to test these hypotheses will be fully described in Chapter 7, Part 3.

It will be recalled that the working hypotheses reflected a focus on (a) classroom behaviour, (b) social interaction, (c) academic functioning and (d) generalization issues (behavioural or response generalization and maintenance in time). As the hypotheses are operationalized here, the following breakdown is apparent: <u>classroom behaviour</u>: hypotheses 1 and 7; <u>social interaction</u>: hypothesis 2; <u>academic functioning</u>: hypotheses 4,5 and 8; <u>generalization:</u> (response generalization): hypotheses 3 and 6; (maintenance in time): hypotheses 1 to 6.

In earlier discussion of means of identifying maladjusted children (Chapter 1) and of assessing change in the context of treatment programmes (Chapter 3), the need for a variety of perspectives or vantage points was stressed. In the present study, at the level of both identification and assessment of change, it was considered important.../

important to seek information from teachers, peers and the children themselves. The measures referred to in the statement of the hypotheses, some of which were employed as identifying and change measures, some solely as change measures, reflect this concern, and afford a variety of perspectives on outcome.

Having described the research hypotheses and the experimental design by means of which they were tested, the next section will commence a description of the manner in which the research study was carried out, beginning with an account of the screen.

(iv) The screen

(a) Location. Selection of the schools to be involved as treatment settings in the study was undertaken in consultation initially with Education Department personnel concerned with the two areas in question (Newcastle-upon-Tyne and Gateshead) and subsequently with the head teachers in the schools concerned. Six schools were selected - three junior high schools in Gateshead and three comprehensive schools in Newcastle-upon-Tyne. A guiding crietrion in the selection process was that the schools be broadly representative of state-run schools in the area. The two cities can be regarded as fairly typical of the industrial conurbations of the north of England, with their associated social and economic problems. The social class distribution of the area is slightly below the national average (Neligan et al., 1974).

(b) The child population. Some 3,339 children were screened in the two consecutive years of screening in the six schools. The numbers of boys and girls were roughly equal (52.6 and 47.4% respectively). The mean age of the group was 11.8 years. Extensive background information was gathered on the study children on factors such as family composition and health, family hazards (e.g. separation, death), social conditions and family relationships and attitudes. In addition, data were gathered on the children themselves by means of parental interview. focusing on behaviour (Kolvin et al., 1975) and temperament (Garside et al., 1975). Information on school environment and characteristics was also sought. These data will not be dealt with at length here. They are presented comprehensively in the main research report (Kolvin et al., 1981) and in Mullin (1979).

It..../

It is useful, however, to consider a few items of background information which might convey something of the important characteristics of the sample of children being dealt with.

First, a high percentage of the children in the six schools received free school meals. These are made available to children from poor or underprivileged homes. In three schools, over 30% of the children were in receipt of free school meals, in one school the figure was under 10% and the others fell between these extremes. Second, nearly one quarter of the families of screen positives were affected by unemployment. 'Unemployed' in this context indicates a breadwinner who had not worked for at least a year. Nearly one-third of these children came from families where the breadwinner's unemployment was a recent or continuing experience, taking those who had been out of work continuously for the previous year or who had not held a job for a continuous year in the preceding three. Third, a marked downward social class gradient, and an under-representation of white collar workers, are evident in these selected cases. In the classification of the breadwinners' occupation according to the Registrar-General's 'Classification of Occupations' (1951), social classes IV and V together, along with the unemployed group, accounted for 49% of the sample.

These data point to the conditions of economic hardship experienced by a substantial number of the children in the study. Difficulties in family relationships were also evident in that some 20% of the children came from families where the parents had divorced or separated. In the 'intact' marriages, some 17% of the partners had separated for a brief period. Single parent families accounted for 17% of the sample, easily exceeding the 10% rate reported for children of all ages (Finer Report, 1974).

Descriptive data derived from screen and additional measures for the children themselves will be presented in the Results Chapter.

(c) Timing. The task of identifying suitable cases was undertaken by a screening procedure (described fully below) applied with all children.../

children entering the first year of senior schooling in the academic year beginning September, 1972, and repeated with the new intake one year later. The project was spread over two years because it was felt that this would lessen the demands being made on school staff through their involvement with the research (e.g. participation in assessment of children and in treatment consultations). Screening was carried out some six weeks into the autumn term. The interval between the start of term and commencement of screening was intended to allow teachers to become familiar with the children, and the children with one another. This was an important consideration, of course, for the meaningfulness of teacher ratings and sociometry.

(d) Obtaining parental consent. Well in advance of the beginning of the screening programme, letters were sent out to parents of all children concerned. This explained that a series of educational assessments would be carried out in the schools as part of a research programme funded by the Department of Education and Science. Parents were invited to contact the schools for further information if they wished. Their consent for inclusion of their child(ren) was to be assumed if they did not indicate otherwise. Very few parents did in fact seek further information and only two children were withdrawn from the screening programme.

(e) Measures employed in screening. On the assumption that multiple measures of adjustment problems were preferable to a single measure, as considered in Chapter 1, measures were sought which reflected the perceptions of the teacher, peers at school, and the child himself. It was felt that the most commonly used screening techniques - teachers' ratings - might be inadequate if not supplemented by other information. Of particular concern was the possibility that teachers might overlook quiet and passive, but potentially troubled children, and that they might/....

might not be sufficiently attuned to the interpersonal difficulties that children might be experiencing. It was felt that sociometric data might supply more accurate information on social functioning and that self-ratings, in complementing teacher- and peer-derived data, might reveal the personal unease and concerns that both other sources of information might bypass.

Three screen measures were chosen: they are briefly described here and fuller details and copies of the instruments are given in Appendix 1.

First, <u>teachers' ratings</u> were sought. The scale employed here was the Rutter Teacher Scale B2 (Rutter, 1967), which was established in usage as a screen instrument (Rutter et al., 1970). This scale can be quickly completed by teachers with a good knowledge of children's day to day behaviour in school. For each of the 26 items, the rater indicates whether or not the behaviour applies on a 3-point scale: doesn't apply (0), applies somewhat (1), certainly applies (2). The items are all overt-behavioural and the rating procedure implies that the teacher will use knowledge of normal children as the standard in judging the child's behaviour. As well as a total score, the scale also yields two subscale scores, for neurotic and antisocial behaviour. The items in the neurotic subscale cover the following behaviours: worrying about many things, being solitary, appearing miserable, having twitches or mannerisms, absence from school for trivial reasons, fearfulness of new situations, complaining of aches or pains, and being tearful on arrival at school. The antisocial subscale covers the following: restlessness, truanting, destructiveness, fighting, bullying, disobedience, inability to settle, lying and stealing.

A second source of information concerned the children's functioning amongst their peers. Direct observation of frequency of interaction, as a means of assessment, was precluded because of the inadequacy of resources to conduct extensive observations. It has also been pointed out that low frequency of interaction is not necessarily an index of maladjustment (Gottman, 1977). Sociometric measures were chosen in preference. Although these measures themselves require further validation, the correlates of low sociometric status reviewed in Chapter One suggest that selected children may well be experiencing genuine difficulties..../

difficulties. Sociometric status may also be an important indicator of long-term adjustment: (Roff et al., 1972). The sociometric instrument was developed for the main study of which the present project comprised a part (Macmillan et al., 1978). It is reproduced in Appendix 1. Each child was asked to choose which three of his classmates he would like to sit beside and which three he would like to play with at playtime and, conversely, those he would not like to sit beside or play with. The total number of choices received by each child was computed for each of these questions. Two scores of interest were derived from this procedure: the first is isolation, which is a low score on the sum of the two positive choices, and the second is rejection, which is a high score on the sum of the two negative choices.

The third screen measure made provision for self-ratings by the children. The measure selected here was the Junior Eysenck Personality Inventory (JEPI) (Eysenck, 1965). This is a sixty-item questionnaire yielding an extroversion-introversion score, a lie score and a neuroticism score. The neuroticism dimension was the one of interest for selection purposes. The high scorer on neuroticism is likely to be chacterised by unnecessary worrying, by feelings of restlessness and anxiety, and by moodiness. These are the characteristics associated with instability, while the stable person (low scorer on neuroticism) is likely to be calm, reliable, carefree and even-tempered.It was felt important to tap such personal feelings in a screen exercise, yet they could be concealed from both teachers and peers. The JEPI offered a useful framework for seeking this information.

(f) Deriving screen identification criteria. It was decided to take extreme scores on each of the screen measures as indicators of deviance. The rationale for the derivation of cut-offs is fully described in Appendix 2. With a multiple criterion screen, one can weight each extreme score equally, or alternatively, assign additional weightings to very high scores on particular measures. With the former system, it is theoretically possible for a child to obtain a very deviant score on one specific measure only, but be excluded because his summed weighted..../

weighted score is not sufficiently high. To avoid this, weighting formulae were adopted which allowed children with markedly deviant teacher- or self-ratings to be selected on that basis alone. The cut-offs and weighting scorers are summarised in Table 5(2).

Table 5(2). Cut-offs and weighting scores on screen criteria

Measure	Cut-off	Weighting score
Rutter B2 Scale		
Total score	9 - 14	1
	15 or above	2
Neurotic subscale (Newcastle modification)	4 or above	1
Antisocial subscale (Newcastle modification)	4 or above	1
Sociometry		
Isolation	0 or 1 positive choice	1
Rejection	12 or more negative choices	1
JEPI Neuroticism	20 - 22	2
	23 - 24	3

The children's scores on each of the screen measures were summed. Those obtaining a total of 3 or more points were regarded as screen positives; those scoring below this total were viewed as screen negatives. From the summary of cut-offs and weighting scores in Table 5(2) it can be seen that children could be selected as screen positives on the basis of extreme scores on either the Rutter B2 scale or JEPI Neuroticism alone, or else by varying combinations of scores from the three instruments. The maximum obtainable weighted score was 9. Of the three screen measures, only the Rutter B2 scale was developed to/....

to discriminate between children who show disorder and those who do not. Even so, all measures at the screen level can indicate only instrumental deviance; whether this is synonymous with difficulties in adjustment is a question for a more detailed level of investigation as described below (section (h)). It is possible, however, that extreme scores on the personality dimension of neuroticism are associated with some degree of malfunctioning or problems of adjustment, as are scores on sociometric criteria which indicate extreme lack of popularity, or rejection.

(g) Carrying out the screen. Assessment sessions for the measures completed by the children (sociometry and JEPI) were organised by the writer in conjunction with a teacher nominated by the head teacher. Administration was conducted in class groups, usually in a single period lasting around 35 minutes. Instructions and questions in the two protocols were all read out to the group so that responses would not be complicated by reading difficulties. The teacher questionnaire was usually completed by the 'class' or 'form' teacher for each group, this being the teacher who was felt to know the children best at this stage of the term, as having had most exposure to them. Collection of screen data took, at most, three weeks. Spare forms were left with class teachers for absentees who missed the group administrations. Returns were not pursued for absentees beyond three weeks.

(h) Checking the efficiency of the screen. In order to check the efficiency of the screen in identifying children with genuine difficulties, all the information available on the selected children and controls - screen data, more detailed school based measures (see below) and data from parent interview - were examined by two psychiatrists. An assessment of diagnosis, severity, and duration of disorder was made.../

made. A detailed analysis of this exercise for the first year's screening has been reported by Macmillan et al., (1980). Only one false positive was recorded, suggesting that the screen had a commendable degree of efficiency in selecting children experiencing genuine problems.

(v) Allocation to regime.

Across the two years' screening, a total of 389 children were selected as 'screen positives' and were available for allocation to the various regimes. With 3,339 children being screened in all, this represents a screen positive rate of 11.7%.

It was not possible to randomize regimes within the two years since the same research staff were involved as consultants/therapists in two of the regimes (PC and GC) and carrying the two regimes simultaneously would have been an excessive burden. Accordingly screen positives in the first year were allocated randomly by school class to either the maladjusted control group or parent counselling/teacher consultation. Their second year counterparts were allocated in the same fashion to behaviour modification or group counselling. This procedure concerned only those classes in which children had been identified by the screen. It was hoped that random allocation by class would help to reduce contamination between regimes within individual schools, but with the inevitable overlap of teachers across classes, this could not, of course, be totally avoided.

As a means of maintaining 'blinds', the names of 'treatment' children were made known only to the staff actively concerned in particular programmes. The names of 'maladjusted' and 'normal' control children were not released at all to staff but were held by the head teacher. Through their involvement in repeated assessments, however, it is likely that their status would have increased in visibility during the life of the project. The age and sex distribution of the children included/.....

included in the two treatment groups and the maladjusted control group are shown in Table 5(3).

Table 5(3). Age and sex distribution of the children

	B.M.	P.C.	G.C.	M.C.
Mean age (in years)	11.8	11.7	11.8	11.8
Age range	11.3-12.2	11.4-12.3	11.2-12.3	11.0-12.3
Number of boys	44	49	33	47
Number of girls	30	34	40	45
Total number	74	83	73	92

From the pool of children not identified by the screen, 124 children were selected at random (half in the first year, half in the second) with the provisos that they be drawn from all six schools, from classes in which screen positives had been recorded, and that the selection should reflect the sex ratio found in the group of screen positives. This group was intended to function as a 'normal' control group, providing descriptive data, for the main study, with which the screen positives could be compared. Another important function was that they were included in all assessments, along with their screen positive counterparts, as a means of attempting to obscure the identity of the latter group as 'problem' children. Data on these normal controls will not be reported here (see Macmillan et al., 1980; Kolvin et al., 1981 for details).

(vi) Organisation of regimes.

(a) Behaviour modification. The introduction of this programme to the schools was staggered to reduce the simultaneous demands that might be made of both the person conducting behavioural observations and/...

and the writer, as the consultant, had all six schools entered the programme at the same time. Accordingly, the six schools were grouped into three pairs. Following identification and allocation of cases, the first pair entered a three week baseline phase. During this phase, 'second level' assessments were conducted, baseline observations were taken on all treatment cases and three teachers' training seminars were held within the schools. As the baseline phase ended for the first pair, and they entered the intervention phase, the second pair entered their baseline period, and so on. The intervention phase lasted for twenty school weeks in each school, taking place in the last two terms of the academic year 1973-74.

(b) Parent counselling/teacher consultation. Following pre-treatment assessments as with the other regimes, this programme was run in the last term of the academic year 1972-73 and the first two terms of the academic year 1973-74. Six social workers were attached to individual schools, undertaking counselling with parents and consultation with relevant teaching staff. The intervention phase lasted for 3 school terms, some 28 school weeks.

(c) Group counselling. Following the assessment of identified children on second-level measures, this programme was run in the third term of the academic year 1973-74. Each group met for ten weekly sessions during the term. The groups were run by six social workers with each worker being attached to an individual school. These were the same six social workers who conducted the parent counselling/teacher consultation regime in the previous academic year, and, in all cases, their school attachment was the same for both regimes.

Full descriptions of these three treatment regimes are provided in the next chapter.

(d) Maladjusted controls. Research staff had no contact with the children in this group other than at the various assessment points. For these assessments, because they were in the same school year as PC children, they would be seen in the same group together. In organising these assessment sessions, the status of the children was not indicated to staff. As mentioned previously, the names of these children were held by the head teacher, and not generally released to the school staff.

Although these children were regarded as untreated controls from the point of view of the study's design, they were not denied any of the customary forms of help that would routinely be available to them, either within the school or from services outside it.

(e) Informing parents and obtaining consent. Parents of all selected children, both treatment and control cases, were visited by social workers on the research staff who were to conduct parental interviews and gather home based data at baseline and the various follow-up points already described. The data gathered from parents is fully described in Kolvin et al. (1981) and will not be dealt with here. Parents' participation was invited in this data-gathering exercise, and the initial visit afforded the social workers the opportunity of explaining to them the nature of the regimes.

Direct parent involvement was, of course, an important component in the parent counselling/teacher consultation regime, so the nature of their participation, for these parents, needed to be fully explained. With regard to the school-based teacher consultation component, parents were told that the discussions with teaching staff would be aimed at assisting their child's progress at school. Parent involvement was not a feature of the other two treatment regimes. In the case of parents of children in the behaviour modification regime, they were told that their children would be part of a research exercise in which variations in teaching procedures would be examined. Parents of..../

of group counselling children were informed that their children would have the opportunity of being involved in discussion groups at school, in which the children would be free to discuss any issues they wished.

In discussion with the parents, it was considered important to avoid implying that their child was being viewed as a special problem at school. Parents were all given the opportunity not to enter the programme if they did not wish to be involved. Table 5(4) gives details of the numbers of cases who did not enter treatment. Cases for whom home data could not be collected were excluded from the study. Dropouts at this stage included those whose parents refused to participate in the first interview with the social workers, and those who withdrew after completing the initial interview. Information on attrition of cases during intervention and follow-up will be presented in the Results Chapter.

The higher number of withdrawals in PC than in any other regime is perhaps not surprising, given the greater demands on parents in this particular regime.

Screen data on dropouts at this stage were compared with the respective data for children who entered treatment. No significant differences were recorded, but the screen data, of course, comprised a limited set of information.

Table 5(4): Withdrawals from study before start of treatment

Regime	MC	PC	BM	GC	Total
Number of children allocated	109	109	85	86	389
Number of withdrawals	17	26	11	13	67
Number of children starting treatment	92	83	74	73	322

In the case of the children themselves, BM children were not informed that they were to be research subjects. This was to avoid making them feel different from others within the class group and to avoid any stigmatisation as 'problem' children. In this regime, procedures were not in fact to be confined to the research cases, although they would be the main focus. PC children were not directly informed by research staff of their status as treatment cases, although from their parents' involvement, they would be aware that school-based discussions would be occurring, designed to assist their functioning in school. GC children were informed by teaching staff of the nature of the discussion groups as settings quite different from the classroom, where they would be free to raise any issues which interested or concerned them.

In the case of control children, the regular participation of parents in the home-based assessments and children in school-based assessment was invited, but neither the parents nor the children were given any information as to their status as 'controls'.

(f) Length of treatment.

Some comment is necessary on the different lengths of treatment across the three regimes. Ideally, treatments should be of equal or similar length, to avoid confounding treatment effects with maturational effects and to equalize treatment exposure. However, the meaningfulness of 'equal' treatment exposure in disparate treatments can be questioned. An example of this problem might be the comparison of intensive daily therapy sessions with another treatment involving once weekly sessions. An important consideration in the present situation was that the length of treatment offered each regime a fair and reasonable test. Practical constraints did also intrude. Thus, in the case of behaviour modification, an extension of the programme into the following school session was considered. However, it was evident that so many of the children in this regime would not have contact with the original group of 'trained' teachers, that a fresh training exercise would be required. This further commitment was not considered to be justified. Accordingly, unlike the PC regime, in which extensive teacher training was not involved, BM did not overlap..../

overlap two academic years.

(vii) Additional measures. Further school-based assessments were conducted with all children included in the study, to provide additional measures of change. These were referred to as 'second-level' measures, to distinguish them from screen instruments. As with the screen measures, brevity and ease of application were important considerations in their selection, so that the demands on teachers, and the time taken away from classroom teaching for group assessments, were kept to a minimum. Measures were sought in the areas of classroom behaviour, school-related attitudes and ability/achievement. They were as follows:

(a) Devereux Elementary Behaviour Rating Scale (Spivack & Swift, 1967).

This is a 44-item rating scale for completion by teachers. It was designed to measure overt behaviours that reflect a child's overall adaptation to the demands of the classroom setting and that may affect his or her achievement in that setting. It yields scores on eleven behavioural factors: classroom disturbance, impatience, disrespect defiance, external blame, achievement anxiety, external reliance, comprehension, inattentive-withdrawn, irrelevant-responsiveness, creative initiative and need for closeness to the teacher. The most attractive feature of this scale for the purposes of the study, was that, in contrast to the Rutter B2 Scale, it provided considerable detail about the child's behaviour in the classroom.

(b) Barker-Lunn Children's Attitude Scales - S7 (Barker-Lunn, undated).

This is a questionnaire for completion by children, containing ten attitude scales: attitude to school, interest in school work, importance of doing well, attitude to class, 'other' image of class, conforming versus non-conforming, relationship with teacher, anxiety in the classroom, social adjustment and academic self-image. These dimensions offered the potential of tapping important aspects of the children's perceptions of school - perceptions which might well be affected by their experiences in a treatment programme.

(c) General Ability Test - (National Foundation for Educational Research).

This test comprises alternate verbal (40 items) and non-verbal (40 items) tasks. The task for each verbal item is to discover the principle or concept underlying a grouping of four words and then to supply the missing word - on a multi-choice basis - to add to a second group of three words, according to the same principle.

The.../

The task is the same for non-verbal items but here the groupings involve shapes. The test yields a verbal ability score, a non-verbal ability score and a total ability score, and in being developed for group administration, it offers a useful and convenient means of assessing children's ability.

(d) Reading Comprehension Test (NFER).

This is a 35-item test of reading comprehension, in which the child is required to choose the correct word from a selection of five given, in order to complete a sentence meaningfully. As a test of comprehension it was felt to offer a more meaningful assessment of 'reading' ability than tests of the reading of single words. Tests of arithmetic/mathematics performance were also sought, but with the absence of a generally accepted measure for this age-group (largely because of the revisions necessitated by decimalisation)none was eventually included.

Fuller details on these four measures are provided in Appendix 3. Home and family based data were also collected on the study children but these are not covered by the present report. They are fully described in Kolvin et al. (1981).

(e) Observational measures in the classroom. In addition to the above measures, direct observations of children and teachers involved in the BM regime were conducted in the classrooms. This form of behavioural assessment was considered an important contribution to the formulation of treatment approaches, and also in its monitoring.

However, it was acknowledged that with limited resources and such a large sample, spread across six schools, it would not be possible to achieve very concentrated coverage of classroom behaviour. Arranging for the observer to see all the children in the treatment regime, in six different schools proved no easy matter. A number of factors created problems for the equal sampling of children. Thus, children or teachers might be absent (often with no prior warning to the observer); some children might be 'available' in double periods, others in singles; the length of period might vary across schools. Along with other difficulties such as timetable clashes, and travelling time from one school to another, these factors led to the adoption of a flexible schedule for observations rather than a rigid and unchangeable one. It became apparent that the demands on the observer to cover.../

cover all the target children in six schools were excessive. Accordingly, one school was omitted from the observational exercise. Baseline data could not be gathered on one child because of absenteeism; fifty-five children were ultimately involved. No classroom observation was conducted with children in other regimes.

The observer, a female psychology graduate, gained experience in the use of the recording system before the beginning of the study, in pilot observations conducted in a school not involved in the main research exercise. She was instructed not to interact with any of the children, and to try to merge into the background as much as possible. The children were informed that she was a student who would be sitting in for a number of sessions to observe teaching methods. To allow the children to get used to her presence, and in anticipation of possible reactive effects, data gathered in her first exposure to each classroom were discarded. In addition, in each classroom period of observation, the first and last five minutes were not recorded since these were often taken up with settling down and transition activities.

The behavioural coding categories employed for children and teachers, and the system of recording adopted, are fully described in Appendix 4. The system of categories is that employed by Becker et al., (1967) with only minor modifications. An interval system of recording was adopted, with a ten-second observe, five-second record arrangement. As described in Chapter 1, such an interval recording system, with behaviours being coded for occurrence or non-occurrence, is useful for dealing with continuous behaviours (e.g. attention, talking, turning round) such as those of concern here. Recording was restricted to the baseline and intervention phases. As the observer completed her contract at the end of the treatment phase, no follow-up data were collected. Even if such measures had been available, they would have posed problems of interpretation since the bulk of the children would by that time have transferred to new teachers in the following session.

The observational data were to be analysed by comparing baseline data with data gathered in the intervention phase. The intervention phase, for this purpose, was not dealt with in a unitary fashion, but was divided into three equal parts, of.../

of 6 weeks and 3 days' duration. This time block was considered a reasonable period in which change might be expected to occur. Data were available for each of these successive phases.

It proved possible to gather a minimum of 24 minutes' observation time for each child in the baseline phase. Any data gathered beyond this were discarded. To ensure that similar periods of observation time were compared in the statistical analysis, 24 minutes was adopted as a target for each child in each of the three treatment phases. This time was spread as evenly as possible by the observer across the time span of each treatment phase.

Percentages were computed for the time 'spent' in each behavioural category as a function of total time observed. The categories of particular concern for purposes of analysis were task-relevant behaviour for children, and approval and disapproval rates for teachers.

It did not prove possible to check observer agreement before the start of the study, and this exercise was conducted during the intervention phase. A second observer who was using the same schedule in a separate project participated in the exercise. The two observers simultaneously rated behaviour in a number of classrooms. This involved observation of 33 children in ten separate recording sessions across the five schools, and a total of eight and a half hours' recording. The procedure for computing observer agreement is described in Appendix 4, as is the method for establishing chance agreement. The level of agreement proved satisfactory, ranging from 80% to 100%, with a mean of 92%, for 'task-relevant' behaviour and acceptably above chance agreement levels. For teacher behaviour, agreement for 'approval' ranged from 89% to 100% with a mean of 98%, and for 'disapproval' from 80% to 100% with a mean of 91%.

(viii) Timing of assessments.

The assessments dealt with here involved the three screen instruments and the three 'second-level' measures. All these measures were available as baseline data for all groups. For the behaviour modification regime (BM) and group counselling (GC) these measures were repeated (a) at the end of treatment - Follow-up 1 (FU1); (B) at a point 11 months after end of treatment - Follow-up 2 (FU2), and (c) 36 months from.../

from the baseline - Follow-up 3 (FU3). For the parent counselling/teacher consultation regime (PC), the measures were repeated (a) at end of treatment (FU1); (b) 10 months after the end of treatment (FU2) and (c) 36 months from the baseline (FU3). The timing of assessments for the maladjusted controls (MC) was identical to that of the PC regime.

There is a slight difference in the FU2 intervals from the end of treatment for the two treatment regimes. This was because the time of school year when assessments became due for BM coincided with a holiday period, so the assessment was delayed by a month. It should also be noted that the time-points for follow-ups indicated here refer to the starting point at which assessments began. With large numbers of children to be assessed and problems of availability and absenteeism, complete data collection might take several weeks.

In the course of these assessments, it became apparent that excessive demands were being made on teaching staff, with the completion of questionnaires (Devereux and Rutter scales) and, often with the same individuals, the scheduling and organising of follow-up group assessments. It was evident at one stage in follow-ups, that in the interests of maintaining co-operation and goodwill, some reduction in demands had to be made. Accordingly the Rutter B2 scale, which gave less direct information on classroom functioning than the Devereux, was omitted from the assessment schedule for BM and GC at the end of treatment.

CHAPTER 6

DESCRIPTION OF THE TREATMENT REGIMES

DESCRIPTION OF THE TREATMENT REGIMES

Having described the overall organisation of the study in the last chapter, the present chapter will present descriptions of the three treatment regimes which were compared: (a) behaviour modification, (b) parent counselling/ teacher consultation and (c) group counselling.

(A) BEHAVIOUR MODIFICATION

(i) The selection of teachers

In approaching the task of selecting teachers who would act as 'mediators' of the treatment procedures, it was clear that, despite the importance of selection, this could not be undertaken by the writer. Having been introduced to the school only at this stage, and having no acquaintance with the staff, it was not possible to make judgements about the suitability of individual teachers. Guidance was therefore sought from the head teacher in each of the six schools. The head was requested to discuss the possibility of their involvement with about six teachers whom (a) he thought might be favourably inclined towards involvement in the treatment programme and (b) who had substantial teaching contact with the classes concerned. A full discussion of the nature of the intervention programme was held with each of the head teachers before they approached their teaching staff.

Following the selection of groups of teachers in each school, discussions were held with each group, in order to explain more fully the procedures to be employed and the role and extent of commitment anticipated for the teachers. Ample opportunity was given for questioning any issues pertaining to the intervention in question or the research exercise as a whole. At the end of this discussion, each teacher was given a copy of a document introducing the characteristics, principles and procedures of behaviour modification in the classroom (Macmillan & Kolvin, 1977). They were then invited to attend a series of three group meetings, at one week intervals, in which the behaviour modification approach would be presented and discussed more.../

more fully. It was recommended that the document be read before the first of these.

While there was no obligation to attend these seminars, the group of teachers who commenced these sessions cannot be regarded as volunteers, since the approach by their head teacher and their subsequent nomination for the initial discussion may well have created a sense of commitment, although this would no doubt vary from person to person. This is an issue of no little importance, given its implications for co-operation and motivation in implementation of the programme.

The group which entered the seminar phase comprised 24 female and 15 male teachers. All but eight had more than five years of teaching experience. Table 6(1) shows the distribution of these teachers and the target children across the six schools.

Table 6(1) : Distribution of target children and teachers across schools

School	Number of children	Number of teachers
A	11	6
B	10	6
C	11	7
D	14	7
E	16	7
F	12	6
Total	74	39

The distribution of selected children across classes in each school is shown in Table 6(2). Also shown here is the number of 'trained' teachers to which that class was exposed. It is evident here that classes varied in the number of trained teachers to whom they were exposed. All classes were exposed to at least two such teachers. Of the nineteen classes, thirteen were exposed to two trained teachers, five classes to three teachers, and one class to four teachers. Table 6(2) also shows the total number of children in each class.

Table 6(2): Distribution of target children and trained teachers across classes

School	Class	Number of Target Children	Class Size	Number of BM teachers
A	1	4	30	2
	2	5	30	3
	3	2	31	3
B	1	2	29	2
	2	3	30	2
	3	2	30	2
	4	3	31	2
C	1	8	28	4
	2	3	30	3
D	1	2	29	2
	2	5	32	2
	3	7	30	3
E	1	4	30	2
	2	2	29	2
	3	2	31	2
	4	8	31	3
F	1	4	30	2
	2	5	29	2
	3	3	30	2

(ii) Training teachers in the application of behavioural procedures

A major consideration in organising the teacher training programme was that the activities should be dovetailed with teachers' other commitments and that disruption of the schools' working timetable be minimised. 'Training' was conceived of as having three components: (a) reading the basic document referred to above, (b) the seminar phase and (c) the consultation phase.

The document was intended to introduce the teachers to the behavioural approach. The fundamental characteristics of the approach were outlined, along with the most important guiding principles, and brief descriptions were given of procedures frequently described in the classroom literature. It is more fully described in Appendix 5. Teachers were expected to have read this document before entering the seminar phase.

In the second stage, three one-hour group meetings were held in consecutive weeks with each set of teachers. These usually occurred either during the lunch-break or after school hours, whichever was more suitable. The time devoted to these seminars was not as much as the writer would have wished, but in discussion beforehand with head teachers and the teachers themselves, it became clear that in view of other commitments and demands on time, a more generous allocation was not possible. For the most part, very full use was made of the one-hour period, and with several groups, discussion spilled over the time allocated. This tended to occur when meetings took place after school hours, but none of the sessions exceeded 80 minutes.

The seminars were devoted to discussion and amplification of the content of the document, description of some relevant studies from the literature, and consideration of how selected cases, from the school in question, might be handled within a behavioural framework. Seminar content is described more fully in Appendix 5. There was a limited amount of role-playing, to demonstrate, for example, the importance of timing teacher attention, the use of combined 'praise-and-ignore' procedures, or the use of soft reprimands.

All the seminars commenced with a presentation of material by the writer but teachers were encouraged to seek clarification or make comments at any time in order to generate a dialogue. The final 15 - 20 minutes were left open for discussion.

In the third phase, consultations with individual teachers were begun, once the group sessions had been completed. All these meetings took place during school hours, either during lunch-breaks or intervals, or during non-teaching periods. In these meetings, discussion was initially geared to the organisation of individual behavioural programmes for the selected children to whom that teacher was exposed. While all available information was drawn upon in the preparation of programmes for individual children, the emphasis in these discussions was always on behavioural functioning and.../

and how the child could be helped by the teacher within the classroom or school context, as opposed to how he might be helped in other settings. An important source of information was the classroom-based observational data supplied by the independent observer.

A set of detailed personal guidelines were worked out for each target child, in collaboration with the teacher. A conscious effort was made to evolve intervention procedures as a collaborative exercise, rather than merely presenting the teacher with a set of prescriptions. The guidelines arrived at were written down in summary form, along with a brief account, in behavioural terms, of the problematic aspects of the child's functioning. Although these guidelines were written down, it was emphasised that they were to be viewed as experimental and flexible, and could be modified in the light of experience and future discussion.

Once treatment strategies had been prepared and implementation begun, the content of the consultation sessions changed to discussion of practical difficulties and progress in the application of particular procedures. These sessions were continued throughout the duration of the programme, lessening in frequency with time, and as teachers were considered to be capably handling the treatment requirements.

'Training' was therefore viewed as a continuing process rather than as a one-shot exercise. The consultation phase was of considerable importance, because it offered some hope of bridging the gap between acquisition of knowledge and actual implementation which the document/seminar phases alone would have been unlikely to achieve. In addition, these individual sessions with the teachers presented the writer with opportunities to encourage and support them in their efforts to apply the procedures.

The average number of consultations conducted per child in the course of the programme was nine (range four - thirteen). The total time taken up by the seminars and the consultations in the six schools was eighty hours, with a mean of 13.3 hours per school (range 11.6 - 15.7 hours). This refers to on-site time only and does not include time spent travelling and preparing material for seminars and consultations.../

consultations or written guidelines.

(iii) Intervention procedures: general considerations

A decision was made at the outset to emphasise social reinforcement procedures - relying upon the systematic and contingent use of teacher attention - rather than upon procedures based on the use of material or concrete rewards. Decisions also had to be made about the extent to which these procedures would be applied. Should they be applied solely with target children or with the whole class? It was felt that the former approach would very likely create adverse responses from the other children who might view such obvious positive selection by the teachers as favouritism. The latter approach, on the other hand, while making the teachers' attention more equitable, would perhaps dilute the intervention since the amount of 'treatment exposure' for the target children would be considerably reduced. Despite this difficulty, this approach was felt to be the more appropriate, so the procedures were to be potentially available to all children in the class, but with special emphasis on the target children,these being the only ones for whom detailed individualized prescriptions were prepared.

Consideration also had to be given to the possible limits of social reinforcement, since it was inevitable that children would vary in their responsiveness and some probably would be unaffected by variations in teacher attention. It was decided that social reinforcement would be pursued with all target children but that other modes of reinforcement would be considered where little success was being encountered. In the event, material and activity rewards were included in the intervention programme for three children who were not judged to be responding to teacher attention alone.

There were several reasons for emphasising social reinforcement. First, difficulties may be encountered in the transition phase in a programme when concrete rewards are phased out. Social reinforcers can be viewed as occurring more 'naturally' in the child's environment than do concrete rewards, and may therefore be more conducive to generalisation and maintenance of behaviour change.

Second, the natural-artificial dimension of proposed rewards is a very significant.../

significant one in an educational context, with continuing debate about the merits of extrinsic rewards and whether or not the emphasis should instead be on 'intrinsic motivation', with children being rewarded for 'learning for its own sake'. While the writer favoured the use of material rewards in certain situations his role could not be one of crusade and conversion, but one that clearly had to accommodate the ideals and philosophies of the schools, represented by head teachers and the various teachers concerned with the programme. There were, in fact, wide differences in attitudes on this issue, some teachers openly favouring concrete rewards and welcoming their applications, others rejecting them out of hand. The decision to minimise the use of concrete rewards was, finally, as much dictated by their artificiality in the regular, natural classroom setting as by objections raised by teachers.

Third, because in the majority of cases very small numbers of 'target' children were being dealt with within a large group (usually around thirty) it was considered invidious to make material rewards openly available to those few children and not to others. Teachers were quick to point out that this would be unfair and possibly lead to counter-productive rivalry and jealousy. In the few cases where material rewards were employed, this was always within the context of a private arrangement between the child, the teacher and the writer. Consideration of cost precluded the extension of such schemes to the entire class. Although the social reinforcement techniques recommended were intended to function as a general classroom-wide management strategy, it was inevitable that slightly more attention was paid to target children than to others by the teachers concerned, since individualised prescriptions were prepared for these selected children. However, this was generally considered to be less obtrusive than administration of material reinforcers to the select few. Some teachers were openly opposed to the notion that certain children should receive extra attention at the expense of others, despite the fact that, even in the absence of a systematic intervention strategy, they tended to give more attention to these children anyway.

Finally, as indicated in the literature review, it was felt that the use of social reinforcement procedures had not been fully explored with children of secondary.../

secondary school age and that an examination of their effectiveness might therefore be worthwhile.

(iv) Defining problems and establishing goals

The behavioural programme could be regarded as having three principal objectives. These involved, first, the modification of social behaviour deemed inappropriate by the teacher in the classroom (e.g. reducing disruptive or noisy behaviour, or excessive talking); second, increasing task-related behaviour (e.g. listening to instructions, improving concentration and persistence with independent tasks), and third, enhancing the interpersonal functioning of children who were isolated and rejected within the classroom group (e.g. by reducing aggressive exchanges or teasing, increasing friendly overtures). These three broad objectives were not of course independent of each other.

In considering interventions for target children, these overall objectives, where relevant, required detailed individual specifications. In considering the needs of a child, the first task was to attempt to describe and define areas requiring change, in observable terms. Thus 'nuisance', 'troublesome', 'unruly', might be translated into 'keeps talking out of turn', 'gets out of his or her seat and wanders around', or 'pinches other children'. This sort of specific definition was seen as essential if there was to be a precise focus for reinforcement Once the behaviour was defined in this way, it was equally important to apply the same stringent criteria in defining goals and objectives for treatment. Targets such as 'well-adjusted' or 'self-actualising' are too vague and nebulous.

Accordingly, in the early phases of consultation, much time was devoted to clarifying concepts of problem behaviour, and encouraging precise initial definitions, clear specification of successive steps through which a child might be expected to progress, and statements about ultimate objectives that were closely tied to observable behaviour. The notion of developing behaviour through successive steps was considered to be crucial to the success of a programme, both in modifying teachers' unwarranted expectations of change, and in conveying the need for gradual 'shaping' of behaviour towards the specified goal.

A most important component in the formulation of an intervention strategy was that.../

that of 'functional analysis'. This involves an attempt to identify features of the child's environment which might be functionally related to the occurrence of his behaviour (Yule, 1977). A relationship is postulated between the problem behaviour and environmental events as a prelude to altering environmental conditions. Thus, teachers were encouraged to identify possible causes and effects of behaviour. For example, what happened before John hit Bill? Was he teased? What were the consequences of hitting? Did Bill cry, or stop teasing? Teachers were encouraged, in particular, to pay attention to their own behaviour vis-à-vis the child, examining ways in which it might elicit or maintain certain behaviours. They were encouraged to vary their own behavour in accordance with the guidelines drawn up, and to monitor its effects. Recognition was given to the obvious constraints on teacher behaviour within the classroom and on the extent to which teachers could freely carry out and monitor such experimental variations.

The identification of problem areas in children's behaviour was approached through discussion with individual teachers, and by consideration of the information supplied by the screening battery and other assessments. It was considered important to place information from these measures within the context of individual teachers' classrooms, since it was recognised that certain kinds of problems might be manifest in certain classes and not others. One of the basic objectives of the programme was to increase the amount of time that children spent 'on-task' or engaged in learning activities as defined by the teacher. This aim was considered a relevant one formost of the target children but the nature of 'on-task' behaviour would, of course, show some variation across the subjects taught. Taking 'on-task' behaviour as one goal, it was apparent that teachers varied in their interpretation of what constituted 'inappropriate' behavbur on this criterion, and discussions soon made clear that tolerance levels and definitions of 'disruptive' behaviour were equally variable. However, the basic categories of behaviour detailed in the observation schedule (see Appendix 4) were discussed with all the teachers as a means of approaching some common basis of understanding. In pursuing the goal of 'on-task' behaviour, teachers were encouraged to provide reinforcement for task attention and studying behaviour, but also for increased/improved output and achievement. Increase in output and achievement..../

achievement may also concomitantly reduce activities which interfere with learning. However, the focus on task attention remained the primary concern, and it was seen as constituting a worthwhile objective in itself. Furthermore, in view of the wide variability in the content of academic tasks across classes and schools, any desired focus on work output or performance would have been difficult to organise in a standardized fashion throughout the treatment group.

The reduction of disruptive behaviour emerged as an important objective. However, the principal strategies recommended for dealing with it were indirect ones, in the sense that the teacher's attention would be channelled towards those behaviours which were incompatible with disruptive behaviour (e.g. working quietly, paying attention, etc) It was hoped in this way to show children that they could only gain the teacher's attention by behaving appropriately. The use of ignoring and soft reprimands in relation to disruptive behaviour is described in the next section. Time-out procedures, which have frequently been employed with such behaviours, were not adopted in the present study. This was because, firstly, most of the schools had a policy which discouraged teachers from sending children out of the classroom for disciplinary purposes and, secondly, the physical facilities (e.g. a small adjacent room) were not conducive to adequate implementation of the procedure. The lack of control over inappropriate reinforcement involved in making a pupil stand outside in the corridor, or sending him to the school office, made such measures quite inadequate as time-out procedures.

In addition to the common goal of increasing task attention, other more individualised problems and objectives were identified where relevant. The screen instruments provided a useful focus for discussion and provided guides to areas of possible difficulty. Thus, the Rutter B2 scale, while not a measure which focuses solely on classroom behaviour, pinpointed many responses which could be modified within the classroom. These included restlessness, difficulty in settling, disobedience, worrying about things, complaining of aches and pains, and unresponsiveness. Reinforcement procedures could be regarded as applicable to all these problems.

The screen measure of JEPI neuroticism could be regarded as even more general

in.../

in its frame of reference than the teacher questionnaire, and as not necessarily focused on the classroom (e.g. 'Do you sometimes feel cheerful and at other times sad without any good reason?' 'Are you touchy about some things?' 'Do you worry about awful things that might happen?') A concern for the behaviour modification programme was that, given the nature of the screen battery, through which children could be selected solely on the basis of a high JEPI neuroticism, some children might be selected for whom classroom-based problems did not exist. This proved less of a difficulty than was anticipated. Seven children were selected by this criterion alone and all of them, to varying degrees, demonstrated difficulties that the teachers felt they could work with in the classroom. These included shyness, excessive sensitivity, worries about marks, achievement, homework, performing in front of the class, volunteering answers, and so on. All these were responses within the 'reach' of a teacher prepared to use, for example, gradual shaping, modelling and prompting procedures to encourage children with such problems, allowing them to experience success, and building up confidence through reaching realistic goals.

The third screen instrument, sociometry, throws up information concerned with success or failure in social interaction, which may be determined to a greater extent by behaviour in the playground than in the classroom. Nevertheless, there were several methods relating to aspects of social interaction which were identified along with the teachers as applicable in the classroom:

(a) shaping procedures for withdrawn children to encourage interaction where classroom activities allowed it;

(b) the creation of special groupings in, for example, project work, craft work, etc., so that desired behaviours could occur (drama activities in two schools created some excellent opportunities);

(c) seating with compatible peers;

(d) reinforcing peers for interaction, or for initiating contact;

(e) with rejected children, reinforcing their co-operative behaviour or friendly approaches, to encourage these to take the place of conduct that usually tended to elicit rejection in others, for example, boasting, teasing.

With all these strategies, opportunities for intervention were to be seized by teachers whenever they presented themselves - in or out of the classroom.

The combination of common and individual objectives as outlined in this section reflected the effort in the programme to confront a spectrum of presenting problems rather than to focus narrowly on one or two limited types of response.

(v) The application of social reinforcement procedures

In the implementation of social reinforcement procedures the writer was guided by the established findings of workers such as Becker et al., (1967), Hall et al., (1968), and Madsen et al., (1968), although their subjects were younger than those selected for our programme. Comparable work with secondary school age children was rare (McAllister et al., 1969). The social reinforcement applied took the form of directed attention, or comments of praise or approval, smiles or a nod, or physical contact such as a hand on the shoulder or a pat on the back, whichever suited the teacher's personal style and preferences. It is quite clear that some people experience marked difficulty in using elements of social reinforcement, whether in normal or therapeutic situations, and with some teachers considerable encouragement and support were required. It was also pointed out that what was reinforcing for any particular child could not be defined a priori - for example, physical contact from a teacher of the opposite sex might embarrass some children; similarly, the class bully might regard the teacher's approval as highly undesirable in view of his or her projected 'hard' image. Accordingly, teachers were encouraged to experiment with different forms of social reinforcement if no success was apparent in the child's response. It was also possible that the reinforcement might vary or wane in effectiveness over time with any one child. For example, the comment 'good girl/boy' might be effective initially, but would soon lose its effect if repeated excessively without variation - so the need for flexibility and innovation in this respect was also stressed.

When.../

When developing specific aspects of behaviour, reinforcement was to be made available contingent upon particular responses. For example, a highly distractible child might be given comments of approval when observed attending to task; a shy, inhibited child might be given comments of encouragement and a smile when he or she made an assertive response in a group situation. It was considered desirable that comments of approval should convey clearly to the child why he or she was receiving such positive attention (for example, 'I'm pleased that you kept on working on these sums, John') rather than being vague, non-specific statements (for example, 'You've been a good boy today.') To encourage teachers to make frequent use of positive social reinforcement they were asked to look actively for specific aspects of behaviour that could be rewarded: to try, in the case of badly behaved children, to 'catch the child being good' (Madsen et al., 1970). This notion underlay the principle of 'shaping' which required teachers not to wait until particular desired responses were fully formed or completed, but to praise the child's approximation to the objective being sought.

Madsen et al.'s notion also encourages teachers to take a positive, guiding stance towards children's behaviour, rather than, as often happens, being trapped into responding - usually with disapproval - to inappropriate behaviour.

Where possible, teachers were to attempt to reinforce behaviour that was incompatible with the conduct they were trying to eliminate. For example, a child given to talking to neighbours might be praised for reading a book quietly, or an aggressive child might be given approval for co-operative responses. For the child who seeks teacher attention, and finds it rewarding, the desired behaviour should theoretically 'displace' its inappropriate counterpart.

Directing comments of approval at a child implicitly conveys information or cues to other children as to how they ought to behave. The child who receives approval may thus function as a 'model' for the others. Teachers were.../

were encouraged to use this 'modelling' opportunity: for example, in situations where a target child was behaving badly, an adjacent child who was behaving well could be praised.

Minor instances of disruptive or deviant behaviour were to be ignored, on the assumption that attention directed to such behaviour, even of a critical, disapproving variety, might tend to reinforce and unintentionally increase it (Madsen et al., 1968). The use of disapproval was not ruled out, however, and was recommended where it was impossible to continue ignoring misbehviour, such as when the learning situation was being disrupted, or when any child was put in physical danger. An important consideration here was the balance achieved between positive and negative controls, and a heavy emphasis on the former was recommended. Where disapproval was employed, soft, private reprimands were considered preferable to the loud, public ones, which may have an adverse effect by creating the very situation that a wrongdoer desires - publicity and a spot in the limelight (O'Leary et al., 1970). Constant loud nagging may also create tension and resentment and serve to perpetuate misbehaviour. In addition to the systematic reinforcement of desired behaviour, teachers were also encouraged to indicate to the children the kinds of behaviours which were considered desirable in his/her classroom, and also those which would not be deemed appropriate. These guidelines and limits could be mentioned from time to time.

The ground rules for applying reinforcement procedures may be summarised in the following list of recommendations to the teachers:

(1) Praise and ignore: praise the behaviours that are to be strengthened and developed and ignore the ones that are to be weakened and eliminated. Vary the reinforcement given.

(2) Avoid disapproval, but if it is necessary, use soft reprimands audible only to the offender, rather than loud ones.

(3) Try to 'catch the child being good': look for behaviour which you can reinforce.

(4)..../

(4) Don't wait for completion of a task or perfect performance - praise attempts or approximations.

(5) Capitalise on modelling effects, e.g. praise a child who is behaving appropriately when the 'target' child is misbehaving.

(6) Remember the balance of your attention: give much more in the way of praise than reprimands in each lesson.

(vi)Applying a behavioural contract system

In three cases, where it was judged that social reinforcement was having no worthwhile effect, and where disruptive behaviour was a major problem, behavioural contracts, giving access to concrete rewards or preferred activities were developed. (It should be pointed out that these three children were not the only ones who proved unresponsive to teacher attention. In other situations where a similar lack of response was observed, the teachers in question did not wish to apply 'rewards' other than those involving attention.) The initial step was to discuss with the child why such a system was being developed. This involved examining the implications of disruptive behaviour for the child's own learning and for the functioning of the class in general. The contract was presented as a means of helping the child to plan his or her own behaviour consciously, and of affording greater control over its consequences. The arrangement was regarded as a private one, between the child, the writer and the particular teachers involved. The confidential nature of the exercise was essential if the administration of concrete rewards on an individual, rather than a whole-class basis, was to be viable.

A list of positive and negative types of behaviour was drawn up in discussion with the child and the teacher, these being tailored to the needs of the individual case. Points values were attached to each of these, with points being earned for positive behaviour, and deducted for negative behaviour. A card bearing these details was given to the child, so that there was no doubt about which types of behaviour would be considered positive and which negative. Teachers awarded or deducted points on the basis..../

basis of behaviour during lessons, informing the child of the details at the end of the lesson. Points totals were also recorded on another card, which the child could keep.

The first points target was always easy to achieve, so that the child could make a good start, but subsequent targets became progressively more difficult, so that back-up rewards became more distant in time and were gradually faded out. Points targets were always decided upon in discussion with the child.

Rewards were administered by either the writer or a teacher, and were made available as soon as conveniently possible after a target had been reached. They usually took the form of items such as pencils, rulers, felt-pen packs, notepads or small toys such as plastic animals or soldiers. Activities were structured as rewards with one child who was allowed extra time in caring for laboratory animals.

(vii) Provisions for maintenance of treatment effects

It is evident from findings surveyed in the literature review that unless specific steps are taken to facilitate and encourage the maintenance of treatment effects, the changes obtained may not persist once the programme has ended. This consideration was one that determined, to a great extent, how the reinforcement programme was develpped and applied, because the most appropriate time for planning for maintenance is not when a programme is over but rather when it is being designed and constructed. The decision to emphasise social reinforcement was perhaps the most important element in this consideration.

In those few cases in which material rewards were given they were always accompanied by comments of approval and praise, and they were gradually phased out towards the end of the programme. In fact, with all forms of reinforcement there was a reduction in frequency and intensity towards the end of the programme. This was guided by previous findings that behaviour may be more effectively maintained by intermittent rather than continuous reinforcement (Bijou & Baer, 1978) and by a desire to lessen the contrast between treatment and post-treatment environmental conditions.

When the consultation programme ended, further written guidelines for follow-up management were made available to the teachers, with written feedback on the children's progress in treatment. Because the end of the consultation programme coincided with the end of the school session, it was inevitable that in the new session, most of the treated children would no longer have teaching contact with any of the teachers involved in the programme. Although this presented a serious difficulty as far as maintenance was concerned, it was unfortunately one that could not be tackled effectively because of lack of resources.

An attempt was made to involve the parents of treated children in group discussions of behavioural management. This was designed to achieve a spread of treatment effects as well as their maintenance.

Apart from encouraging the transfer of any gains obtained in the school setting, parental involvement was also considered to be of potential value in providing continuity of reinforcement against the background of frequent change and inconsistency of management in the school. Given the demands of the school-based programmes, the parental exercise could not be undertaken concurrently, so its feasibility was not explored until near the end of the consultation phase. Lack of resources also prevented a personal approach or home visit to the parents to explain the nature of their possible involvement and seek their co-operation. The approach was made by letter, with pre-paid reply cards but, although initial non-responders were sent a second card, only 12 per cent of the parents showed interest, so the project was reluctantly abandoned.

(B) PARENT COUNSELLING/TEACHER CONSULTATION

Six professionally qualified social workers were the central research staff in this regime. One was attached to each of the six schools in the project, and also, concurrently, to one of six primary schools involved in a similar treatment exercise. The present report is not concerned with this latter involvement. The programme lasted for three school terms: the third term of session 1972-73 and the first two terms of session 1973-74.

(i) Training of social workers

The social workers had not worked within schools prior to this study, so a 3-month period of training was organised. This had a number of components:

(a) weekly seminars in the Newcastle-upon-Tyne University's Child Psychiatric Unit dealing with casework and other psychotherapeutic topics, with special emphasis on psychodynamic procedures applicable within schools;

(b) additional seminars, geared to school-based work, provided by a school social worker, a psychologist and a head teacher of a school for the mal-adjusted;

(c) casework supervision sessions, dealing with a small caseload, provided by two senior caseworkers and two consultant psychiatrists;

(d) attendance at sensitivity groups.

Informing the training approach was a concern with the effective transfer of social work skills to the school setting and the diplomacy with which the entry and functioning of the social worker would have to be handled.

Consideration was given to relevant aspects of consultation theory, relationship development skills (Anderson, 1974), 'threat reduction techniques' (Klein, 1959) and the notion of the school as a social institution with its own philosophies, and structures of organisation and management.

(ii) Teacher consultation

The aims of consultation in this part of the programme were:

(a) to heighten teachers' awareness of the psychodynamic aspects of pupil behaviour and to use this as a basis for child management;

(b) to provide teachers with relevant family and social information which would assist them in formulating a management plan.

The early stages were regarded as a settling-in period for the social workers. During this time, the social worker was spending in the order of one day per week in the school, with discussions with teachers being organised around both the social worker's own additional commitments, such as family visits and administrative duties, and also the teachers' timetable..../

timetable. Discussions were usually scheduled during break, lunch and non-teaching periods.

It was important at the outset to establish a trusting relationship with teaching staff before embarking on discussions concerning the children's difficulties. Early exchanges therefore dealt with issues to do with the concept of social workers functioning alongside teachers, possible clashing of roles, and the feasible distribution of teachers' activities across 'pastoral' and purely 'educational' spheres.

A limiting factor in the programme was the number of teachers (up to 15) with whom a single child might have contact in the course of a week, making for potential inconsistency in caring and management. The teachers, on the other hand, might be exposed to between 200-300 children each week, in large classes, placing severe limits on the extent to which they could get to know individual children well. The solution adopted was to concentrate discussions on teachers who taught a child for more than three lessons per week. In practice, social workers collaborated with, on average, 16 teachers in order to discuss all the identified children in that school. Within such a group, the official pastoral staff proved to be key figures and much of the individual programming for children was tailored around the work of these staff. A consistent feature of interaction within the programme was that the majority (80%) of contacts were initiated by social workers. Two-thirds of all discussions lasted a minimum of ten minutes and an average of four such consultations took place each week. Half as many discussions (35%) were classified as briefer contacts. Group consultation, though offering an economical use of time, proved difficult to arrange because teachers were not available at the same time.

Detailed discussions with at least one teacher, or briefer discussions with several teachers, were held on each weekly set of school visits for 24% of the children. Some 52% were discussed on a similar basis once every three weeks and 24% less than once a month. Thus, the amount of detailed contact varied considerably.

(iii) Method of consultation

Four strands can be identified in the consultation approach adopted. First, social workers could provide emotional support for teachers. Second, they could act as a 'sounding board' to enable teachers to formulate and crystallise ideas about management. Third, they could engage teachers in discussion of child management which encouraged an appreciation of children's inner feelings rather than a dwelling on superficial explanations of their behaviour (Long et al., 1971). Within this framework, the child's disturbed behaviour could be examined more carefully for what it might reveal about the child's mechanisms for coping with stress, his conflicts and preoccupations. Fourth, the social workers encouraged links between home and school both by passing on of information relevant to treatment plans, in either direction, or by encouraging teacher-parent contacts.

Treatment plans were conceptualised around six main headings, which functioned as guidelines for the social workers as to where the main emphasis in consultation should be put:

(a) individualising the curriculum, so that, as far as possible, each child's emotional, social and academic needs could be met within the school system. Teachers' awareness of children's stress or sensitivity in certain areas, for example, could be heightened, extra help for underachievers could be arranged, or areas of success might be pursued for children lacking in self-confidence;

(b) variation of classroom activities and structure, to meet the needs of certain children. For example, curricular restructuring might call for interspersing academic work with free and exploratory activities; physical re-structuring might involve re-arrangement of seating, perhaps to bring a child closer to the teacher for support, or to minimise distraction by other children;

(c) discussion of classroom management and rules, so that exploration was allowed of issues such as consistency across teachers, differing views about discipline and the possible resulting confusion and frustration for children;

(d) linking home and school, as already mentioned;

(e) discussion of child's home environment, in order to increase understanding of the child's behaviour in school. For example, disciplinary patterns at home might help explain responses to authority figures in school, or family stresses or crises might underlie poor performance educationally, or patterns of absenteeism or truancy;

(f) discussion of extra-curricular activities, so that benefits to children's adjustment could be explored through non-academic media, such as sports, hobbies, etc.

(iv) Parent Counselling

The work undertaken in this part of the programme was modelled on the short-term task-centred approach (Reid & Epstein, 1972). The basic aim of the work with parents was to promote in them an awareness of the way family factors might influence the child's behaviour and performance, particularly within the school. This involved (a) providing parents with detailed information about relevant aspects of the child's functioning in school, (b) promoting parental support for changes being sought at school and (c) providing direct social work help for attendant family problems, either by task-centred approaches (with problems of interpersonal conflict, for example) or by traditional methods (with physical health or financial problems).

The initial stage involved the assessment of problems and planning of casework with the parents. Early sessions also served to add to the baseline information that was already available, to establish rapport, to explain programme objectives and to gain some understanding of the interactive factors at work in the family. Discussion was also initiated on school-based problems identified by the parents and the social worker and how family problems might impinge on the child's school functioning. After two introductory visits, the number and severity of problem areas, and treatment aims, were recorded in consultation with the back up team of social work tutors and psychiatrists. Many families had multiple problems but.../

but with the brief, focussed approach being adopted, it was impracticable to try to convey help with all the presenting difficulties. Four areas were chosen for further intervention in the majority of cases: emotional problems, marital problems, parent-child relationship problems and problems about schooling.

The programme in total consisted of up to ten visits per family, including the preliminary sessions, but the number of visits varied, with most families receiving 4-6 visits in all. In only a minority of visits were both parents seen together, although in 55% of the cases, there was at least one joint interview.

(C) GROUP COUNSELLING

This programme was run in the second year of the project. The principal research staff involved here were the same six social workers who conducted the parent-counselling/teacher-consultation regime in the first year. As with their earlier involvement, each worker was attached to one of the six schools in the project, and also to one of six primary schools to conduct a related play therapy regime. This report is not concerned with the primary school exercise.

(i) Rationale

The groups were presented to the children as settings in which they would be able to discuss anything they wanted. The focus of discussion was always on the 'here-and-now' interaction in the group. The therapist did not attempt to channel discussion in particular directions, assuming a non-directive stance in keeping with the method of Carl Rogers (1952). His therapy is based on the assumption that maladjustment is the result of attempting to preserve the existing self-concept from the threat of experiences which are inconsistent with it, leading to selective perception and distortion or denial of experience by incorrectly interpreting those experiences. If certain conditions are provided in therapy, the incongruence between self and experience can be overcome.

Within the guiding framework of Rogerian principles, the therapists, functioning .../

functioning as a kind of sounding-board for the children, sought to create those conditions identified by Rogers, in which interpersonal encounters could serve to deepen their understanding of their own feelings and their relationship to others. The conditions considered necessary and sufficient for therapeutic change are (1) empathy, (2) positive regard and (3) genuineness (Rogers, 1959). These are usually regarded as therapist attitudes rather than techniques per se. The first condition of empathy requires that the therapist focusses on the client's phenomenal world but not in an attempt to interpret or diagnose. The assumption is made that the experience of being understood promotes personal growth. One way that the therapists could convey empathy was to reflect or mirror what the children said by paraphrasing their remarks in an understanding way. Such reflective techniques give feedback that the child can clarify or expand on. The second condition is unconditional positive regard. The therapists could demonstrate this attitude primarily by being non-judgemental, and showing acceptance and concern for a child as he is, with his capacities and limitations, strengths and weaknesses. The third condition of genuineness involves the therapist being 'real' in the relationship. This involves therapists expressing their own personal response to children's feelings in a consistent way. It also involves responding to children in need not in terms of their status, role or in terms of some stereotype, but naturally and spontaneously.

(ii) <u>The structure of the regime</u>

There were seventeen discussion groups in all, consisting of four or five children of the same sex. The decision to work with same-sex groups was taken in the light of suggestions from the literature that sexual anxieties in the 11-12 year age group could significantly inhibit interaction in mixed sex groups.

Each group met for ten sessions over one school term, meeting on a weekly basis. Sessions ran for the duration of one lesson period (thirty to sixty minutes). All of the groups were conducted on school premises, two.../

two of them in large empty classrooms, the others in appropriately sized rooms. The physical structure emphasised the nature of the group as a group, with chairs arranged in a circle.

(iii) The therapist training programme

In addition to the training experiences to which the therapists had been exposed for the purposes of the parent-counselling/teacher consultation programme, further training was provided for the present regime. This involved: (a) pilot group experience in which each therapist conducted both a playgroup and a group therapy programme in schools not involved in the main project, each for ten sessions;

(b) training sessions with a psychotherapist experienced in child psychotherapy;

(c) sensitivity groups, in which the therapists were in the role of group members. These sessions were also attended by the psychotherapist. Their purpose was seen as one of providing the opportunity for the trainees to gain insight into their own feelings and responses, the effect they have on others, and the functioning of groups through personal experience;

(d) opportunities for consultation with the child psychiatrist who was supervising the group counselling project.

During the treatment programme itself, each therapist was allocated a half-to one-hour personal supervision session each week with the child psychiatrist, and there were also opportunities each week to discuss special problems with the psychotherapist.

(iv) Therapy processes

Most of the groups at some time raised the issue of confidentiality of discussion material. It was decided that the groups' discussions would be strictly confidential and that the therapists would not function as a bridge between the children and their teachers. However, some limited discussions took place between therapists and teachers, centring on the child and his or her problems rather than on the content of group sessions. It .../

It was emphasised with teaching staff that in the interests of group members' feelings of trust and security, such confidentiality was necessary. Therapist-teacher contacts were seen as helpful in allaying potential anxieties or antagonisms that might arise. Pastoral staff were the ones most involved here.

In running the groups, it was important to reach agreement on what constituted acceptable behaviour in the group, so limits on certain kinds of behaviour needed to be established. These were not spelled out to the children in advance, to avoid creating a negative atmosphere or the expectation of trouble, but were made clear as the situation demanded. The principle guidelines covered: (a) any infringement of the general school rules, such as smoking, climbing out of windows, damaging school property, etc., (b) any behaviour that seriously disrupted group interaction such as wandering about or leaving the room and (c) any overt physical aggression shown to either other children or the therapist.

The total mean attendance rate in the groups fell during the programme, starting at a mean of 87.2 per cent and ending with a mean of 71.4 per cent. As the programme took place in the summer term, some decline in attendance might be expected as a general trend towards the end of term. Checking of attendance registers showed that children were absent from school as a whole and not just from the group sessions.

Discussion content showed considerable variation across sessions. In early sessions, the focus tended to be on school issues, such as likes and dislikes of teachers, peers, school rules and discipline and so on. The next dominant focus tended to be home related matters, but there was much variation between groups in the extent to which such themes were explored. In one boys' group home life was not touched on at all. Leisure activities also had a prominent place in the early sessions.

After the first three or four sessions, discussions tended to move away from topics outside the group and to focus more on interaction between group members. This was accompanied by an increase in the expression of..../

of negative affect, both verbal and non-verbal, and a reduction in the frequency of prolonged silences. There was also a marked increase in the extent to which limits had to be imposed. The changes observed seemed to correspond with the first two of the three stages in group functioning described by Yalom (1975): (a) the initial orientation stage with hesitant participation and a search for meaning, (b) the second stage of conflict, dominance and rebellion and (c) the third stage of increasing cohesiveness and working through.

A number of dynamic group themes were identifiable during the course of the programme. These included monopolisation of the group by one member who dominates discussion; assumption of the therapist role by individual group members; conflict about self disclosure, especially in the earlier stages; splitting the group to avoid anxiety (e.g. one member moved her chair away from the therapist and started up a private conversation); scapegoating, with certain members taking the brunt of aggressive feelings; problems of limit-setting, often arising through children's attempts at avoiding personal conflict or anxiety by setting up confrontations with the therapist.

The children were informed at an early stage that the group would run for ten sessions, and the subject was reintroduced in later sessions. Some members dealt with their feelings in a very direct way by questioning what the therapist would go on to do the following term, whether she would be continuing to work in the school and so on; others showed sadness that the sessions were coming to an end.

Chapters 5 and 6 have described the experimental procedures employed in the study, and the three treatment regimes. The following chapters, in Part 3, will present and discuss the results of the study.

PART 3

DATA ANALYSIS AND DISCUSSION OF RESULTS

CHAPTER 7

RESULTS

RESULTS

This chapter will present the outcomes in relation to each of the experimental hypotheses. Detailed interpretations of the findings will be offered in the next chapter. Before presenting the results, the methods of analysis will be described, first, in relation to those hypotheses which concern between-group comparisons (hypotheses 1-6), and, second, in relation to analyses of findings solely within the behaviour modification regime (hypotheses 7 and 8).

(i) Methods of Analysis (Between-Group)

(a) Statistical procedures

Where a hypothesis called for comparison of regimes in relation to improvement, analysis of covariance was employed. While analysis of variance is a legitimate procedure to employ where subjects are randomly allocated to regimes, it does not control for differences between regimes - despite randomization - which may affect outcome. Pre-treatment levels of disturbance, for example, may be correlated with improvement, and analysis of covariance allows the necessary adjustment to be made to final scores to take initial differences into account.

Analysis of covariance shares with the analysis of variance the assumptions of normality of distribution of scores, and of homogeneity of variance. There is evidence, however, that the analysis of covariance is robust with respect to violation of these assumptions (Box, 1953). Only one variable was dropped from analysis in the present set of comparisons because of violation of either of these assumptions. This was 'rejection', one of the two sociometric variables, and it was omitted from analysis because of its J-shaped distribution of scores.

An additional assumption of analysis of covariance is that within-group regressions are homogeneous. In other words, it is assumed that the regression coefficients of the variate (improvement in the present analyses) upon the covariates (e.g. initial score) are the same for each regime. Tests for homogeneity of regression were conducted for each variable across regimes, although Winer (1971) indicates that analysis of covariance is robust with respect to this assumption.

None.../

None of the measures involved in testing the present hypotheses, was excluded on the basis of this inspection. Details are provided in Appendix 6.

In addition to these formal assumptions governing the use of covariance, it is also important to conduct the analysis with measures of adequate reliability. Lord (1960), for example, showed that covariance adjustments with fallible measures can result in relationships that are actually opposite in direction to the true underlying relationship - a clearly undesirable outcome. Some workers are uneasy in having to deal with adjusted change scores, on the grounds that they lose a 'feel' for the data, as if the 'real' changes were somehow being by-passed. While reference to the raw data is not, of course,precluded, the risks of making errors by calculating improvement on unadjusted data by far outweigh any benefits associated with 'feeling close to the data'.

Where the present research hypotheses carried a prediction of maintenance of improvements in time across follow-ups, three change scores for each variable were computed. Thus, typically, for each variable involved in testing such a hypothesis, change scores would be computed for the differences between baseline and end of treatment follow up (FU1) values, baseline and midline follow-up (FU2) values, and baseline and final follow-up (FU3) values. (Unless otherwise stated, hypotheses 1-6 below involve covariance analyses based on these three adjusted change scores for each variable.) The change means were all standardized, being based on change scores having a standard deviation of unity at the final follow-up. For initial to end of treatment and initial to midline follow-ups change means were adjusted to make them comparable with those at the final follow-up.

Change means pertaining to all variables relevant to hypotheses 1-6 were adjusted by analysis of covariance, taking into account four covariates. These comprised initial level, general severity of maladjustment, non-verbal ability score and an index of social functioning within the family. These were found to be the most important covariates amongst a larger number examined, and of these four, initial level proved to be the most influential. These covariates can be briefly described. 'Initial level' was the pre-treatment value of the variable in.../

in question. 'General severity' was based on a clinical judgement by a psychiatrist, drawing on all information available on a child. Each child was rated on a four point scale: (a) no disturbance, (b) slightly disturbed, (c) moderately disturbed and (d) markedly disturbed. 'Non verbal ability score' was derived from the General Ability Test. The 'index of social functioning within the family' was a composite score derived from indices of 'psychosocial hazards' and 'social risks' identified in parental interviews. 'Psychosocial hazards' consisted of ratings on the following: significant separations from mother or father in the first five years of life, parental loss, child(ren) in care, parental physical and mental illness. For 'social risks', ratings were made on overcrowding in the home, housing conditions, adequacy of toilet facilities, unemployment, and social agency contacts.

Where the probability of the F ratio was equal to or less than 0.05 (i.e. where differences between the regimes were unlikely to have occurred by chance), paired comparisons were carried out using analysis of covariance. The statement of the hypotheses implies only a comparison of behaviour modification with each of the other groups. However, in order to make more meaningful statements about the performance of behaviour modification in relation to that of the other groups, this limited set of comparisons was not adequate. For example, these comparisons might show significant differences in favour of behaviour modification, but it is important to know whether these advantages are specific to it, or shared with one or other treatment regime. Comparisons of all groups were therefore required. The comparison procedure adopted was the Newman-Keuls (Winer, 1971). This technique takes account of the fact that where more than two means are arranged in order of magnitude, the probability of two means differing is altered by putting them in order.

In determining whether the statistical tests supported the experimental hypotheses, conventional values for the level of significance, $p < .05$ and $p < .01$, were adopted. Even though the hypotheses were written in a directional form, thereby permitting the use of one-tailed tests, two-tailed tests were employed consistently, since these offered a more stringent test of the hypotheses.

There are two important issues that need to be considered in connection with statistical significance. First, where a large number of statistical tests are conducted, the attainment of one or two significant results may simply be a function of chance. For example, if 35 comparative tests are conducted, and two are found to be at the .05 level of significance, such an outcome would be so close to a chance occurrence that little weight should be attached to the results. Where such a possibility exists in relation to the present analyses, it will be mentioned in connection with the hypothesis in question.

A second issue concerns the practical or clinical significance of findings that are of clear statistical significance. This question has already been given some consideration in the earlier discussion of methodological issues. Strictly speaking, all that can be concluded from statistically significant findings is that they are not very likely to be due to chance. Whether the results obtained have any clinical usefulness cannot be determined by the statistical tests alone. In this connection, tables of raw scores are provided for variables on which significant outcomes were recorded (Tables 7(10) to 7(14)). These provide a guide to the average changes within groups.

(b) Subject losses and missing data

Over the time span of the study - three years from baseline to final follow-up - data losses occurred in a number of ways. This could occur with cases withdrawing from treatment, with families moving out of the area, or because of difficulties in obtaining certain items of data. With children transferring schools, even well outside the area, it was sometimes possible to obtain self-completed and teacher-completed information. However, it was not considered reasonable to ask a new school to conduct sociometry for a single child so this information was not usually pursued in the case of transfers. The changed context would, of course, render such data of questionable value anyway. Other data losses, rather than case losses, occurred with inability in some situations to obtain teacher ratings and self-completion data with children who were persistent absentees or unpredictable in their attendance.

The effect of such losses was to reduce the pool of cases with complete data.

As.../

As one might anticipate, more problems were encountered with missing items of data than with missing cases. The loss of cases or data across follow-ups creates the possibility that the analysis of outcome is conducted with a biased sample. This was checked in analyses for the main research project by comparing all school and home data available at baseline for cases which subsequently dropped out with the data for those cases who remained in the study. Perhaps surprisingly, no differences were found between the two groups in these comparisons (Kolvin et al., 1981).

At least partial information was available for 274 cases, at the final follow-up, comprising 85% of the total original sample at the point of beginning treatment. For individual regimes the respective percentages were: behaviour modification: 88%; group counselling: 78%; parent counselling/teacher consultation: 88% and maladjusted controls: 86% Group counselling therefore emerges with the highest attrition rate.

One of the consequences of missing data is that the statistical analysis can focus on all the data available at a particular follow-up point, or on the common group of subjects for whom complete information is available across all follow-up points. In a comparison of these two methods of analysis, in fact, the differences were found to be trivial, and did not affect the attainment of statistical significance (Kolvin et al., 1981).

At FU1, a common group was defined by taking into account both school-based measures and the home based measures gathered for the purpose of the main study (Kolvin et al., 1981). By FU2 and FU3, however, with home based data having proved the more difficult to gather, the common group was defined with reference to school-based measures alone. Hence the numbers presented for FU1 differ from those presented for FU2/FU3. In addition, for the purposes of the main study, factor analyses were conducted as a means of organising the grouping of data. The Devereux and Barker-Lunn scales were analysed separately from the other school-based measures in these factor analyses, so the numbers of cases in the outcome analyses presented here for these two measures differ slightly from those for the remaining body of measures. No cases are included in any of the FU2/FU3 analyses which do not appear in FU1 analyses.

(ii) Methods of Analysis (Within-Group)

(a) Statistical procedures

Two of the hypotheses (7 and 8) involved examination of data pertaining only to the behaviour modification regime. Hypothesis 7, firstly, predicted increases in task-related behaviour during the intervention phase. Rather than comparing baseline data with 'treatment' data viewed in a global fashion, an attempt was made to detect changes within the treatment period itself. Thus, the treatment period was divided into three equal time phases. The observations available for each of these phases were averaged to provide one value for task-relevant behaviour for each phase.

A comparison of baseline task-relevant behaviour values and the three subsequent treatment values (phases 1,2 and 3) was carried out by analysis of variance for repeated measures. When the analysis of variance proved significant, the question of which means differed significantly was examined, using the Studentized Range Statistic (Winer, 1971).

Hypothesis 8 predicted that increases in task-relevant behaviour within the behaviour modification regime would be associated with gains in reading comprehension as assessed at end of treatment. The hypothesis was tested by computing a product-moment correlation between reading comprehension change scores (end of treatment score minus baseline score) and task-relevant behaviour change scores (final treatment phase values minus baseline values).

(b) Subject losses and missing data

The analyses for both these hypotheses were conducted on 55 cases, this being the number left following the omission of one school from the observational exercise and the exclusion of one irregular attender for whom complete observational information was not available.

(iii) Outcome

(a) Between-group analyses

This section will present the results in relation to each of the hypotheses concerned with comparisons between groups. To avoid congestion in the text, a number.../

number of tables and figures detailing these results are located at the end of this chapter. These include details of adjusted change means (Tables 7(6) to 7(9)); figures which illustrate these changes across follow-ups (Figs. 7(1) to 7(7)), and changes expressed as raw scores for those variables for which statistically significant results were found (Tables 7(10) to 7(14)). Tables giving details of the statistically significant differences obtained in paired comparisons of regimes will be located in the text.

Hypothesis 1: There will be significantly better improvement in the behaviour modification (BM) regime than in maladjusted controls (MC), parent counselling/ teacher consultation (PC) and group counselling (GC) on teacher ratings of behaviour in school and classroom at (a) end of treatment, (b) midline and (c) final follow-up assessments.

The analyses here focused on adjusted change scores, firstly, for the eleven Devereux factors, for baseline to FU1, baseline to FU2, and baseline to FU3, and, secondly for the three Rutter B2 scores (Total, Neurotic and Anti-Social) for baseline to FU2 and baseline to FU3 only (Rutter B2 ratings were not collected at the end of treatment.) Table 7(1) below, shows the statistically significant differences that emerged from all the paired comparisons conducted. Other data relating to this hypothesis are located at the end of the chapter. They are: adjusted change means in Tables 7(6) and 7(7) and Figs. 7(1) and 7(2), and raw scores in Tables 7(10) and 7(11).

At the end of treatment, BM had significantly higher scores than all other groups - MC, PC and GC - on the Devereux Factor 10: "creative initiative". In addition, GC had significantly higher scores than PC. In describing "creative initiative", Spivack & Swift (1967) suggest that "scores on this factor are positively related to achievement, measuring the degree to which the child exhibits personal involvement in, and positive motivation to contribute to, the classroom learning situation. A high score reflects a child who tends to become constructively and personally involved and actively initiates behaviour and ideas which are relevant to the classroom work." (p.17)

On the Devereux Factor 11: "Need for closeness to the teacher", BM and GC had.../

TABLE 7(1). SIGNFICANT DIFFERENCES (PAIRED COMPARISONS): TEACHER RATINGS

Item		F ratio	p <
	(a) base to end of treatment		
Devereux Factor 10 (creative initiative)	BM>MC	11.35	.01
	BM>PC	19.38	.01
	BM>GC	6.13	.05
	GC>PC	4.95	.05
Devereux Factor 11 (need for closeness)	BM>MC	21.80	.01
	BM>PC	22.26	.01
	GC>MC	15.26	.01
	GC>PC	17.23	.01
	(b) base to final follow-up		
Rutter B2 Total score	BM>MC	6.25	.05
	GC>MC	14.61	.01
Neurotic score	BM>MC	6.47	.05
	GC>MC	7.06	.05
Anti-social score	GC>MC	9.85	.05

Note: no significant differences were recorded at base to midline follow-up;

> means 'better than' in this table.

had significantly higher scores than MC and PC. This factor is described as tapping the extent to which children like to be close to, seek out and offer to do things for the teacher. High scores suggest that the teacher is positively valued by the child. Scores on this factor have been found to be related to achievement in special class settings and with primary age children. None of the comparisons on the other Devereux factors reached statistical significance.

For comparisons conducted at the midline follow-up, no significant differences emerged between regimes on either the Devereux or the Rutter scales. At the final follow -up , no significant differences were recorded on the Devereux factors. On the Rutter B2 scale, BM and GC had significantly lower scores than MC on both the Total score and the Neurotic subscale score. In addition, GC had significantly lower scores than MC on the Anti-social subscale. The Neurotic subscale includes items such as: worrying about many things, fearfulness of new situations, appearing miserable, and complaining of aches and pains. Items on the Anti-Social subscale include: restlessness and inability to settle, destructive and disobedient behaviour, fighting, bullying, lying and stealing.

Raw scores shown in Tables 7(10) and 7(11) indicate that MC showed no improvement at all on these teacher ratings, and no gains were evident for PC on the two Devereux factors considered here.

In summary, therefore, there is little support for Hypothesis 1, in that, while some positive changes were observed for BM, they were not, with the sole exception of "creative initiative", specifically in favour of BM, being shared with GC. No Devereux scores reached significance following the end of treatment assessment, and Rutter scores did not reach significance till the final follow-up. There is therefore no maintenance of changes in the sense that gains on the same measure are evident in consecutive follow-ups. GC also had one advantage (Rutter Neurotic at final follow-up) over MC which was not shared by BM. Finally, with BM being involved in 99 comparisons in all on Devereux measures (33 at each follow-up) the five significant results obtained are close to the level of chance findings.

Hypothesis 2:.../

Hypothesis 2: There will be significantly better improvement in the BM regime than in MC, PC and GC on sociometric measures at (a) end of treatment, (b) midline and (c) final follow-up assessments.

Since the Rejection measure was excluded as unsuitable for covariance analysis, the measure focused on here was Isolation. Children's "Isolation" score reflects the number of positive choices they have received from peers choosing classmates to (a) sit beside and (b) play with. The significant differences recorded in paired comparisons are shown in Table 7(2) below. At the end of the chapter, adjusted change means are shown in Table 7(8) and Fig. 7(3), with raw scores in Table 7(12).

TABLE 7(2) SIGNIFICANT DIFFERENCES (PAIRED COMPARISONS): SOCIOMETRIC SCORES

Item		F ratio	p <
	(a) base to end of treatment		
Isolation	BM>MC	7.47	.05
	BM>PC	12.11	.01
	BM>GC	7.40	.05
	(b) base to midline follow-up		
Isolation	BM>GC	8.69	.05
	(c) base to final follow-up		
Isolation	GC>MC	5.82	.05
	GC>PC	7.15	.05

Note: > means 'better than' in this table.

At end of treatment, BM did significantly better than all other groups - MC, PC and GC. By the midline follow-up, this superiority was maintained only against GC - this was the only significant comparison. At the final follow-up no differences in favour of BM were recorded. However, at this stage GC did significantly better than MC and PC.

The evidence for hypothesis (2) is therefore mixed. At the end of treatment, there is clear support. By midline, BM had lost its advantage over MC and PC, retaining it only in relation to GC. At the final follow-up, this advantage had dissipated and at this stage it was GC that was doing better than both MC and PC.

Hypothesis (3): There will be significantly better improvement in the BM regime than in MC, PC and GC on personality self-ratings at (a) end of treatment, (b) midline and.../

and (c) final follow-up assessments.

The measure analysed here was the neuroticism dimension on the JEPI. This scale includes items which tap unnecessary worrying, moodiness, and feelings of restlessness and anxiety. Table 7(3) below shows the significant differences in paired comparisons. Adjusted change means are shown in Table 7(8) and Fig. 7(3) with raw scores in Table 7(12).

TABLE 7(3): SIGNIFICANT DIFFERENCES (PAIRED COMPARISONS): PERSONALITY SCORES

Item		F ratio	p
	(a) base to end of treatment		
Neuroticism (JEPI)	BM>PC	10.75	.01
	(b) base to midline follow-up		
Neuroticism (JEPI)	BM>MC	12.30	.01
	BM>PC	32.67	.01
	MC>PC	3.92	.05
	GC>PC	15.95	.01
	GC>MC	4.77	.05

Note: > means 'better than' in this table; no significant differences were recorded at base to final follow-up.

At the end of treatment, BM showed a significant reduction in neuroticism relative to PC only. At midline, both BM and GC scores were significantly lower than those of both MC and PC. In addition, MC showed a significant improvement over PC. No significant differences were observed at the final follow-up.

Support for this hypothesis is therefore minimal. The only advantage over MC is shared with GC and the only evidence of maintenance is relative to PC. All these advantages were lost by the final follow-up. Table 7(12) shows that while BM neuroticism scores continued to fall, improvement by MC and PC meant that these regimes 'caught up' by the final follow-up.

Hypothesis 4: There will be significantly better improvement in the BM regime than in MC, PC and GC in measures of ability at (a) end of treatment, (b) midline and (c) final follow-up assessments.

This hypothesis was tested with verbal, non-verbal and total ability scores on the..../

the General Ability Test. Significant differences in paired comparisons are shown in Table 7(4). Adjusted change means are shown at the end of the chapter in Table 7(8) and Fig. 7(4), with raw scores in Table 7(13).

TABLE 7(4): SIGNIFICANT DIFFERENCES (PAIRED COMPARISONS): ABILITY AND READING COMPREHENSION SCORES

Item		F Ratio	p <
	(a) base to midline follow-up		
Verbal ability	BM>MC	14.74	.01
	BM>PC	18.71	.01
	GC>MC	10.60	.01
	GC>PC	16.30	.01
Non-verbal ability	BM>MC	5.86	.05
	BM>PC	6.90	.05
	GC>MC	8.21	.05
	GC>PC	11.33	.01
Total ability	BM>MC	11.45	.01
	BM>PC	14.73	.01
	GC>MC	10.77	.01
	GC>PC	17.55	.01
	(b) base to final follow-up		
Reading Comprehension	BM>PC	10.47	.01

Note:>means 'better than' in this table; no significant differences were recorded at base to end of treatment.

No significant differences were recorded at either the end of treatment or final follow-up assessments. At the midline assessment BM and GC showed significant improvement over MC and PC on verbal, non-verbal and total ability scores.

Support for this hypothesis is very limited in that the gains made by BM were shared with GC. There was no evidence of maintenance.

Hypothesis 5: There will be significantly better improvement in the BM regime than in MC, PC and GC in reading comprehension at (a) end of treatment, (b) midline and (c) final follow-up assessments.

The relevant measure here is the score on the Reading Comprehension Test. Table 7(4) shows the significant differences in paired comparisons. Adjusted change means.../

means are shown in Table 7(8) and Fig. 7(5) at the end of the chapter, with raw scores in Table 7(13).

No significant differences were observed until the final follow-up. At this stage BM showed significant gains by comparison with PC. This was the only significant result.

Again, support for this hypothesis is minimal with the only advantage for BM accruing at the final follow-up, and then only relative to PC. No gains were made which could be tested for maintenance.

Hypothesis 6: There will be significantly better improvements in the BM regime than in MC, PC and GC in school-related attitudes at (a) end of treatment, (b) midline and (c) final follow-up assessments.

The relevant variables here are scores on the ten attitude scales of the Barker-Lunn Attitude to School Questionnaire S-7. These scales are: (1) Attitude to school; (2) Interest in school work; (3) Importance of doing well; (4) Attitude to class; (5) 'Other' image of class; (6) Conforming versus non-conforming; (7) Relationship with teacher; (8) Anxiety about school work; (9) Social adjustment and (10) Academic self-image.

Significant differences in paired comparisons are shown in Table 7(5) below. At the end of the chapter, adjusted change means are shown in Table 7(9) and Figs. 7(6) and 7(7), with raw scores in Table 7(14).

At the end of treatment, BM showed a significant improvement on Scale 1 (Attitude to School) relative to MC. In addition, GC improved significantly on Scale 1 relative to MC. The 'Attitude to School' scale is composed of items concerned with general rather than specific aspects of school (e.g. 'school is fun', 'school is boring', 'going to school is a waste of time', 'I like school').

On Scale 2 ('Interest in school work'), BM showed significant gains over PC. This scale is concerned with both school work in general and particular lessons (e.g. 'school lessons are boring', 'I enjoy most school work', 'I like doing hard sums').

At midline, both BM and GC showed significant reductions on Scale 8 ('Anxiety in..../

TABLE 7(5). SIGNIFICANT DIFFERENCES (PAIRED COMPARISONS): BARKER LUNN SCHOOL ATTITUDE SCALE

Item	F ratio		p <
	(a) base to end of treatment		
Scale 1 'Attitude to school'	BM>MC GC>MC	7.44 7.54	.05 .05
Scale 2 'Interest in school work'	BM>PC	7.53	.05
	(b) base to midline follow-up		
Scale 8 'Anxiety in the classroom	BM>MC GC>MC	8.27 7.45	.05 .05
Scale 9 'Social adjustment'	GC>MC BM>PC GC>PC	5.77 7.34 8.54	.05 .05 .05
	(c) base to final follow-up		
Scale 6 'Conforming versus non-conforming'	BM>PC	9.67	.05

Note: > means 'better than' in this table.

in the classroom') in comparison with MC. This scale is concerned with the child's anxieties, fears, and worries in the classroom (e.g. 'school work worries me', 'I feel scared when the teacher asks me questions about my work', 'it would bother me if I got my work wrong').

Also at midline, GC made significant improvements on Scale 9 ('Social adjustment') in comparison with MC and PC, while BM showed a significant advantage relative to PC only. This scale reflects the child's ability to 'get on' with other pupils in the class (e.g. 'I think the other children in my class like me', 'I have no-one to play with at playtime', 'I have no friends I like very much in my class').

At the final follow-up, the only significant comparison was on Scale 6 ('Conforming versus non-conforming pupil') with BM showing increased conformity in comparison with PC. This scale reflects attitudes to staying out of trouble or mischief, fooling about or creating noise in class. A reduction of support for these attitudes and behaviours is considered here as 'improvement'.

Hypothesis 6 is not supported in that, while BM shows significant advantages at all assessment points, they do not reflect comprehensive superiority over the other groups. Where advantage unique to BM is evident ('Interest in school work' and 'Conforming versus non-conforming pupil') this is only in relation to PC. None of the gains observed was maintained across assessments.

(b) Within-group analyses

Hypothesis 7: Observational ratings of classroom behaviour in the BM regime will show an increase in task-related activities during the period of intervention.

The mean percentages for task-relevant behaviour for baseline, and phases 1, 2 and 3 of the treatment period were 77.5%, 82.7%, 79.5% and 84.1% respectively as shown in Fig. 7(8). The analysis of variance conducted on these four phases proved significant ($F = 4.28$, df 3/162, $p < .01$). Paired comparisons employing the Studentized Range Statistic q showed that the first and third treatment phases differed significantly (both at $p < .05$) from the baseline. The second phase did not, nor did treatment phases differ from one another.

The hypothesis is borne out, although the improvement during the treatment period was not consistent. In addition, it should be pointed out that the evidence for a true.../

true treatment effect is weakened by the absence of control data.

Changes in observational data relating to teacher behaviour as opposed to child behaviour are not included in the present hypotheses. Teacher behaviour is considered in Appendix 7.

Hypothesis 8: Increases in task-related behaviour within the BM regime will be associated with gains in reading comprehension as assessed at the end of treatment.

The product-moment correlation between the two change scores proved not to be significant ($r = -0.06$, $N = 53$, $p = 0.33$), so this hypothesis was not supported.

This chapter has briefly presented the outcome in relation to each of the experimental hypotheses. A full interpretation of the results will be presented in the following chapter.

TABLE 7(6). ADJUSTED CHANGE MEANS: TEACHER RATINGS (1)

adjusted change means

Item	MC	PC	BM	GC	F ratio	p
Rutter B2	(a) base to midline follow-up					
Total score	0.16	0.35	0.30	0.47	1.04	.38
Neurotic score	0.21	0.22	0.25	0.35	0.21	.89
Anti-social score	0.12	0.33	0.20	0.43	1.82	.14
n =	72	72	62	51		
	(b) base to final follow-up					
Total score	-0.07	0.20	0.42	0.60	4.54	< .01
Neutoric score	-0.01	-0.04	0.37	0.33	3.82	< .05
Anti-social score	-0.02	0.22	0.17	0.46	3.12	< .05
n =	72	72	62	51		

TABLE 7(7) ADJUSTED CHANGE MEANS: TEACHER RATINGS (2)

Item		MC	PC	BM	GC	F ratio	p
Devereux		(a) base to end of treatment					
Factor	1	0.07	0.15	0.09	0.02	0.41	.75
"	2	0.02	0.17	0.12	0.14	0.61	.61
"	3	-0.22	0.00	-0.02	0.09	2.09	.10
"	4	-0.02	0.13	0.08	0.09	1.19	.32
"	5	0.07	0.26	0.11	0.14	0.84	.47
"	6	0.23	0.30	0.33	0.12	0.89	.45
"	7	0.03	-0.01	0.14	0.05	0.43	.73
"	8	0.24	0.17	0.33	0.16	0.65	.59
"	9	-0.06	0.19	0.16	0.04	1.68	.17
"	10	-0.03	-0.28	0.38	0.08	8.19	<.05
"	11	0.32	0.36	-0.24	-0.15	12.67	<.05
	n =	75	78	70	63		
		(b) base to midline follow-up					
Factor	1	0.09	0.18	0.07	-0.10	1.40	.26
"	2	0.03	0.19	0.07	0.06	0.56	.65
"	3	-0.20	0.02	-0.09	0.00	1.07	.37
"	4	0.01	0.15	0.11	0.20	0.75	.52
"	5	0.11	0.23	0.26	0.14	0.49	.69
"	6	0.26	0.29	0.19	0.24	0.17	.92
"	7	0.02	-0.04	0.13	-0.04	0.49	.69
"	8	0.23	0.18	0.36	0.21	0.58	.63
"	9	0.00	0.22	0.17	-0.16	2.59	.06
"	10	-0.05	-0.28	-0.06	0.12	2.34	.07
"	11	0.35	0.32	0.40	0.04	2.09	.10
	n =	72	69	62	54		
		(c) base to final follow-up					
Factor	1	0.02	0.05	0.08	0.10	0.12	.95
"	2	0.08	0.11	0.01	0.22	0.59	.62
"	3	-0.39	-0.34	-0.32	0.00	1.43	.23
"	4	-0.15	-0.08	-0.19	-0.01	0.36	.78
"	5	0.15	0.10	0.09	0.05	0.13	.94
"	6	0.08	0.16	0.40	0.03	1.96	.12
"	7	-0.06	0.06	0.28	0.21	1.66	.18
"	8	0.03	0.29	0.37	0.30	1.88	.13
"	9	-0.17	-0.12	0.12	0.07	1.19	.32
"	10	0.04	0.15	0.29	0.27	0.87	.46
"	11	0.47	0.29	0.31	0.08	2.11	.10
	n =	72	69	62	54		

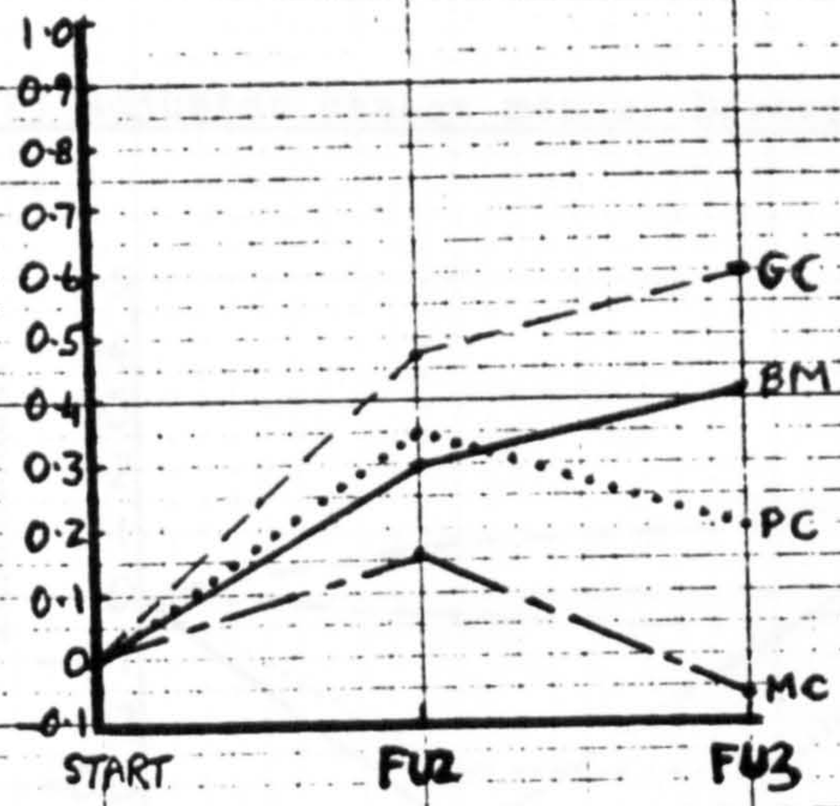
TABLE 7(8) ADJUSTED CHANGE MEANS: SOCIOMETRIC, PERSONALITY, ABILITY AND READING COMPREHENSION MEASURES

	adjusted change means					
	MC	PC	BM	GC	F ratio	p
(a) base to end of treatment						
Isolation	-0.06	-0.18	0.39	-0.04	5.47	<.01
Neuroticism	0.23	-0.03	0.49	0.22	4.22	<.01
Verbal ability	-0.01	0.04	0.10	0.09	0.67	.57
Non-verbal ability	-0.05	0.01	-0.05	0.04	0.28	.84
Total ability	-0.02	0.02	0.03	0.08	0.37	.78
Reading comprehension	0.29	0.21	0.39	0.30	0.78	.51
n =	81	80	68	61		
(b) base to midline follow-up						
Isolation	-0.04	-0.12	0.15	-0.31	2.93	<.05
Neuroticism	0.23	-0.04	0.80	0.66	11.00	<.01
Verbal ability	-0.02	0.05	0.47	0.47	9.21	<.01
Non-verbal ability	-0.08	-0.01	0.31	0.37	4.79	<.01
Total ability	-0.04	0.02	0.41	0.45	7.89	<.01
Reading comprehension	0.30	0.25	0.40	0.39	0.52	.67
n =	72	72	62	51		
(c) base to final follow-up						
Isolation	-0.21	-0.29	0.05	0.19	3.62	<.05
Neuroticism	0.51	0.69	0.89	0.62	1.64	.18
Verbal ability	0.64	0.56	0.68	0.86	0.80	.50
Non-verbal ability	0.56	0.56	0.60	0.80	0.79	.50
Total ability	0.60	0.57	0.69	0.91	1.60	.19
Reading comprehension	0.74	0.53	1.02	0.72	3.05	<.05
n =	72	72	62	51		

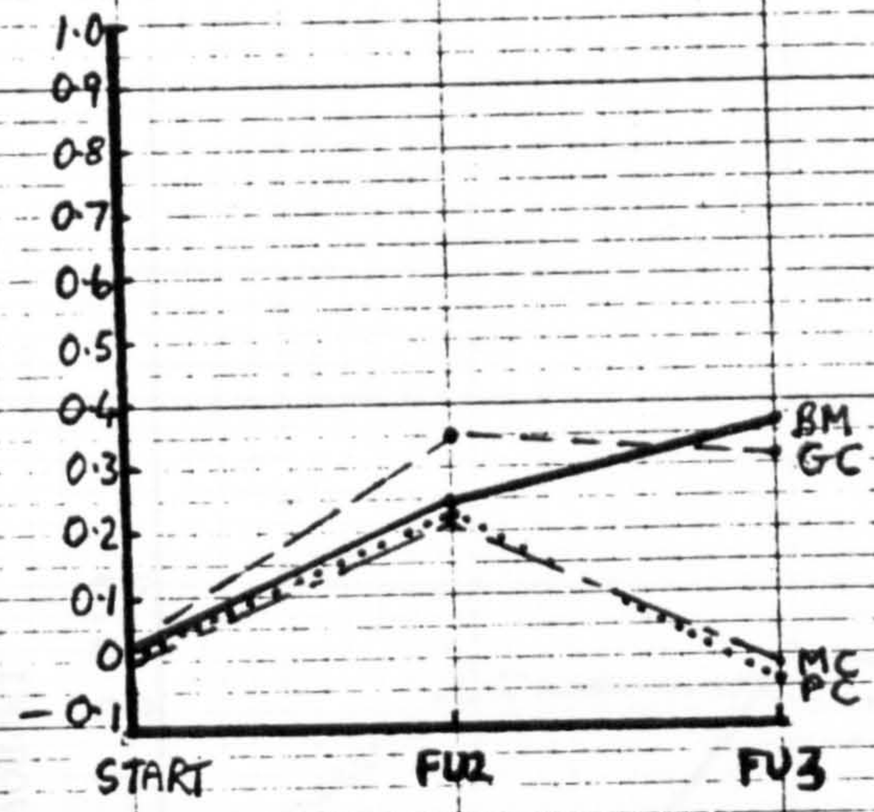
TABLE 7(9). ADJUSTED CHANGE MEANS: ATTITUDE MEASURES

	adjusted change means					
Barker-Lunn	MC	PC	BM	GC	F ratio	p
	(a) base to end of treatment					
Scale 1	-0.44	-0.33	0.01	0.01	4.71	<.01
" 2	-0.38	-0.45	-0.03	-0.15	3.31	<.05
" 3	0.01	-0.31	-0.07	-0.24	1.81	.15
" 4	0.00	-0.14	-0.08	-0.12	0.34	.80
" 5	-0.02	0.11	0.16	0.06	0.44	.72
" 6	-0.04	-0.28	0.02	-0.16	1.92	.13
" 7	-0.28	-0.28	0.04	-0.20	2.10	.10
" 8	-0.05	0.02	0.23	0.13	1.24	.30
" 9	-0.08	-0.10	0.09	0.17	1.45	.23
" 10	0.08	0.00	0.30	0.05	1.91	.13
n =	82	77	68	60		
	(b) base to midline follow-up					
Scale 1	0.43	0.39	0.21	0.12	1.69	.17
" 2	0.34	0.45	0.31	0.28	0.42	.74
" 3	-0.05	0.29	0.11	0.12	1.43	.24
" 4	0.00	0.19	-0.01	-0.03	0.78	.51
" 5	0.02	-0.06	0.18	-0.09	0.82	.48
" 6	0.03	0.21	0.12	0.08	0.64	.59
" 7	0.29	0.29	0.17	1.11	0.56	.64
" 8	0.06	-0.04	-0.37	-0.32	3.83	<.05
" 9	0.05	0.11	-0.26	-0.32	3.67	<.05
" 10	-0.04	0.02	-0.23	-0.18	1.08	.36
n =	79	73	65	57		
	(c) base to final follow-up					
Scale 1	0.37	0.35	0.14	0.03	2.03	.11
" 2	0.22	0.10	0.09	0.22	0.54	.78
" 3	0.10	-0.02	0.10	-0.11	0.56	.64
" 4	0.16	0.06	0.12	0.20	0.25	.86
" 5	-0.18	-0.32	-0.42	-0.22	0.81	.49
" 6	0.08	0.39	0.00	0.12	3.22	<.05
" 7	0.42	0.31	0.10	0.01	2.85	<.05
" 8	-0.36	-0.30	-0.67	-0.33	1.89	.13
" 9	-0.30	-0.23	-0.37	-0.28	0.30	.83
" 10	-0.04	-0.11	-0.34	-0.08	1.11	.34
n =	79	73	65	57		

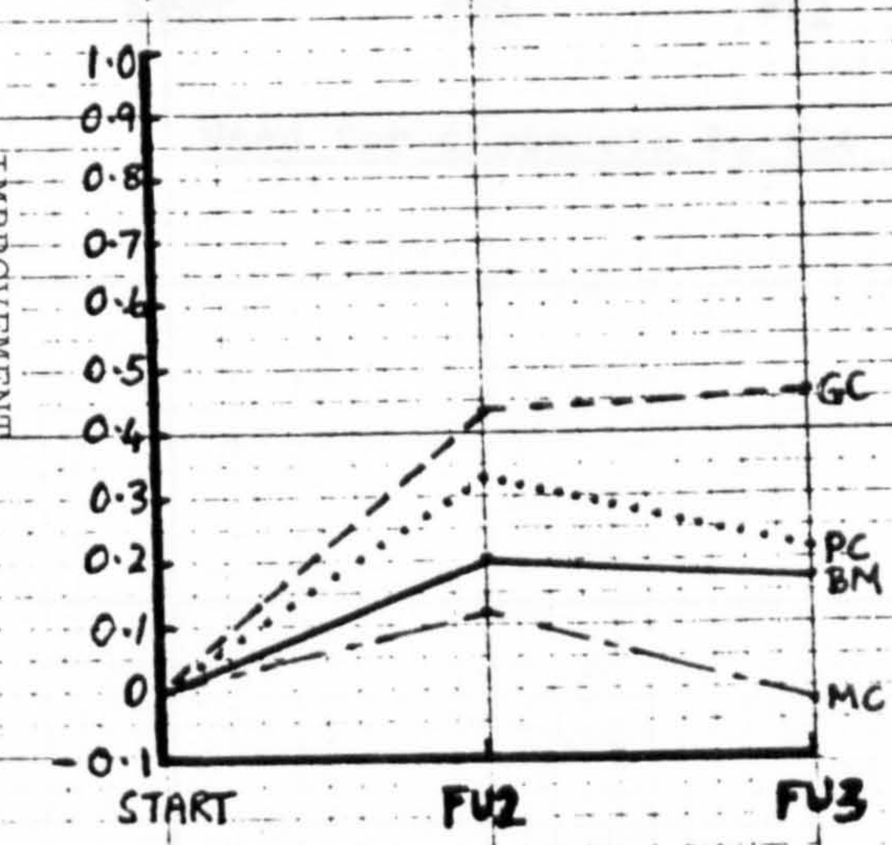
FIGURE 7 (1) Adjusted Change Means: Rutter B2 Scale



TOTAL SCORE

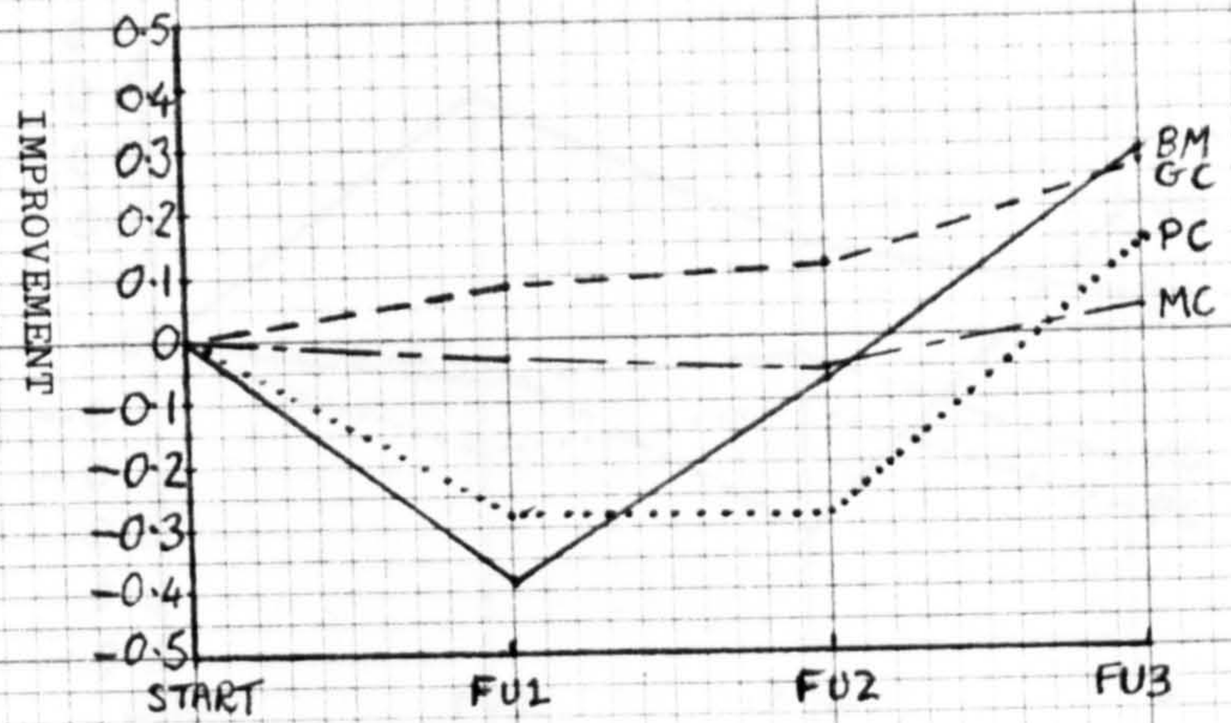


NEUROTIC SCORE

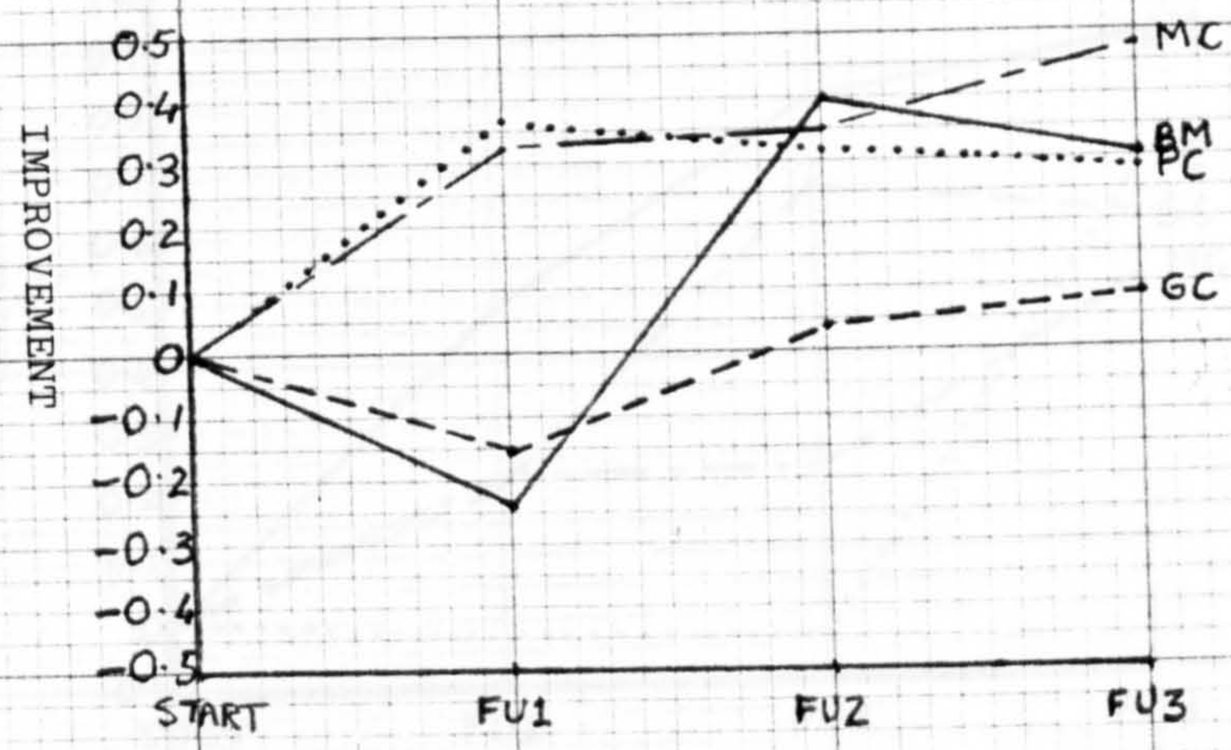


ANTI-SOCIAL BEHAVIOUR SCORE

FIGURE 7(2) Adjusted change means: Devereux scale

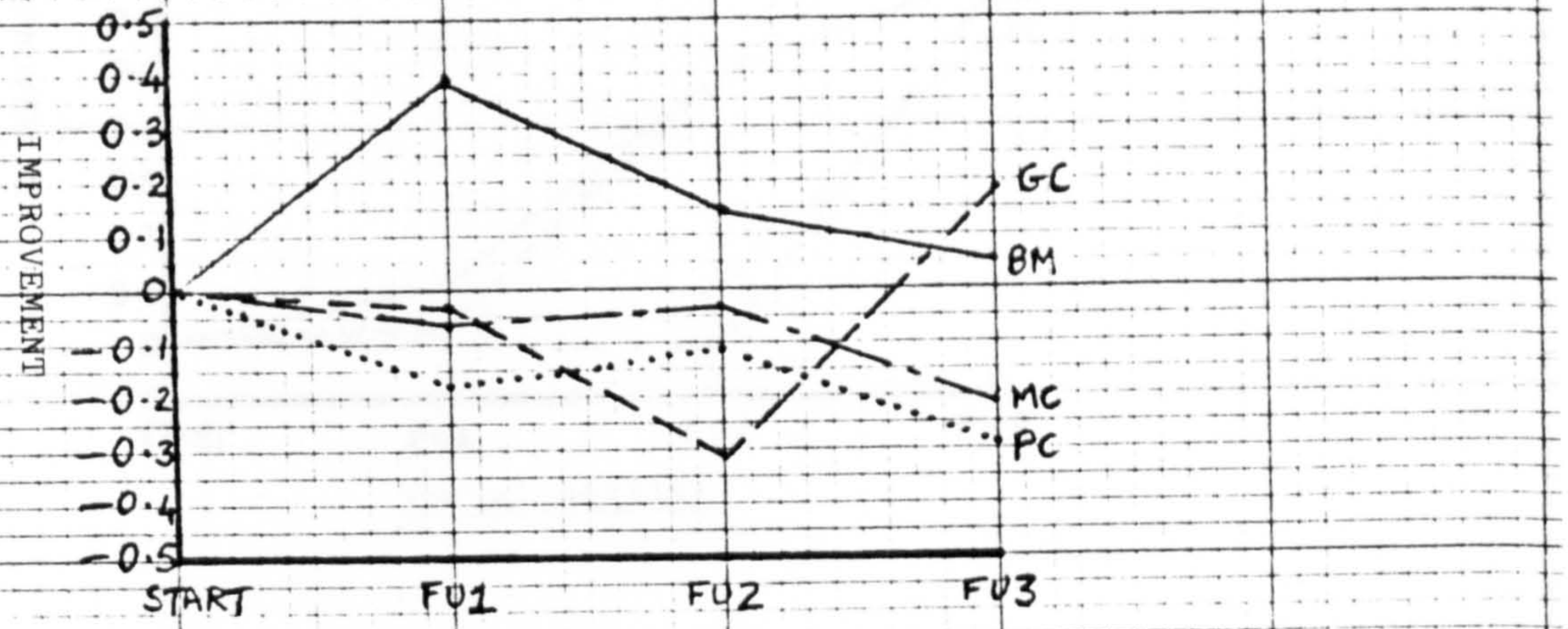


Creative initiative

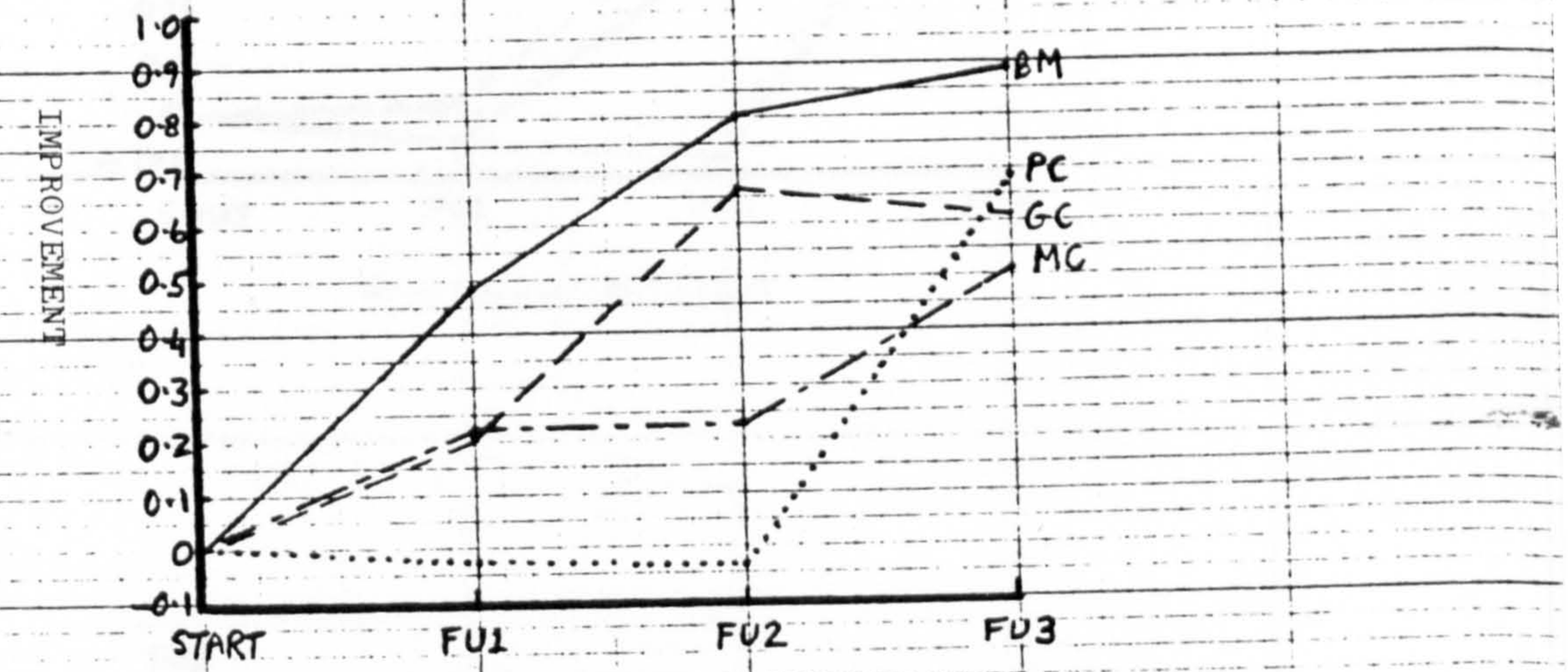


Need for closeness to the teacher

FIGURE 7(3) Adjusted change means: isolation (sociometry) and neuroticism (JEPI)

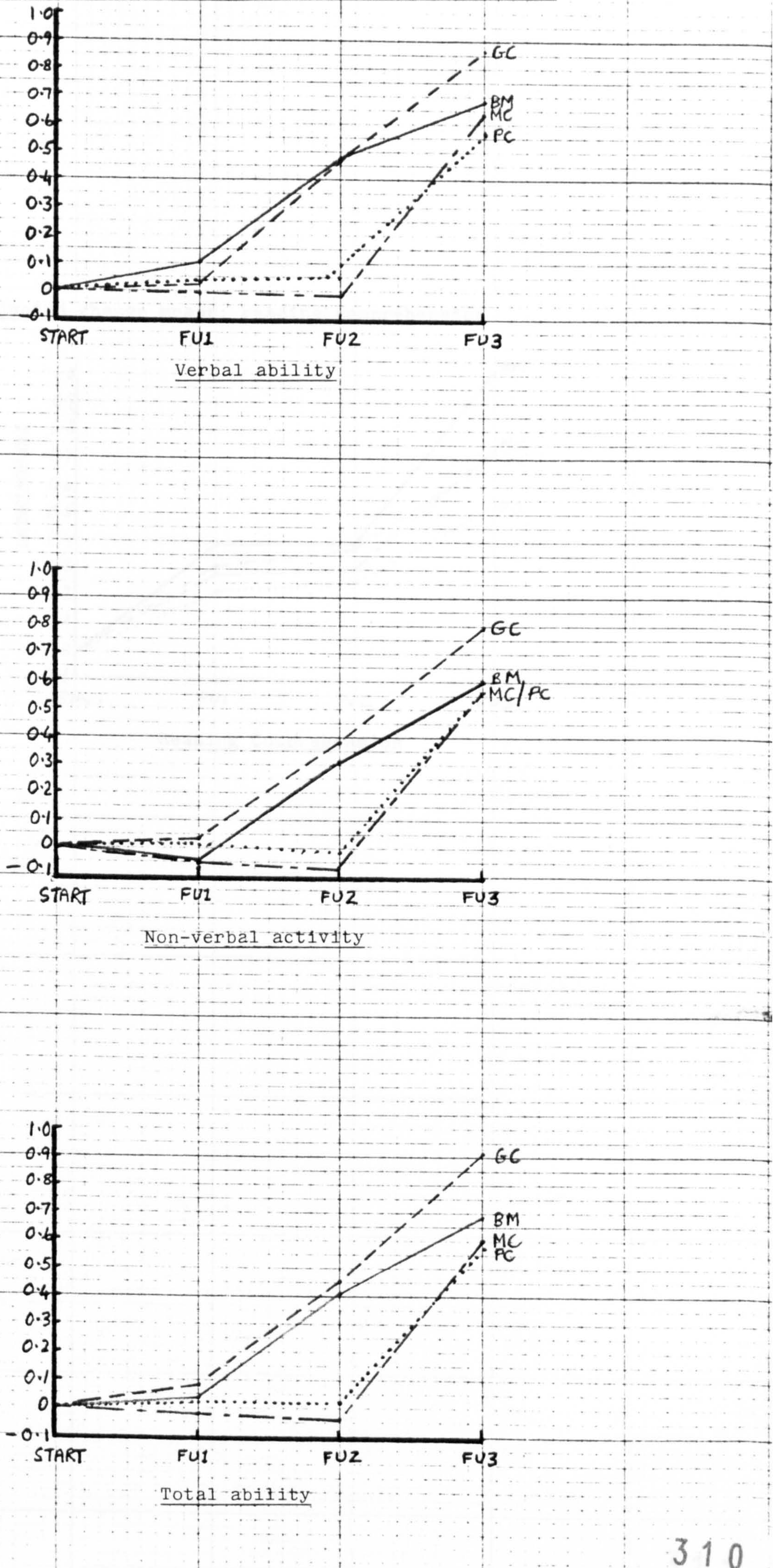


Isolation



Neuroticism

Figure 7(4): Adjusted change means: ability measures

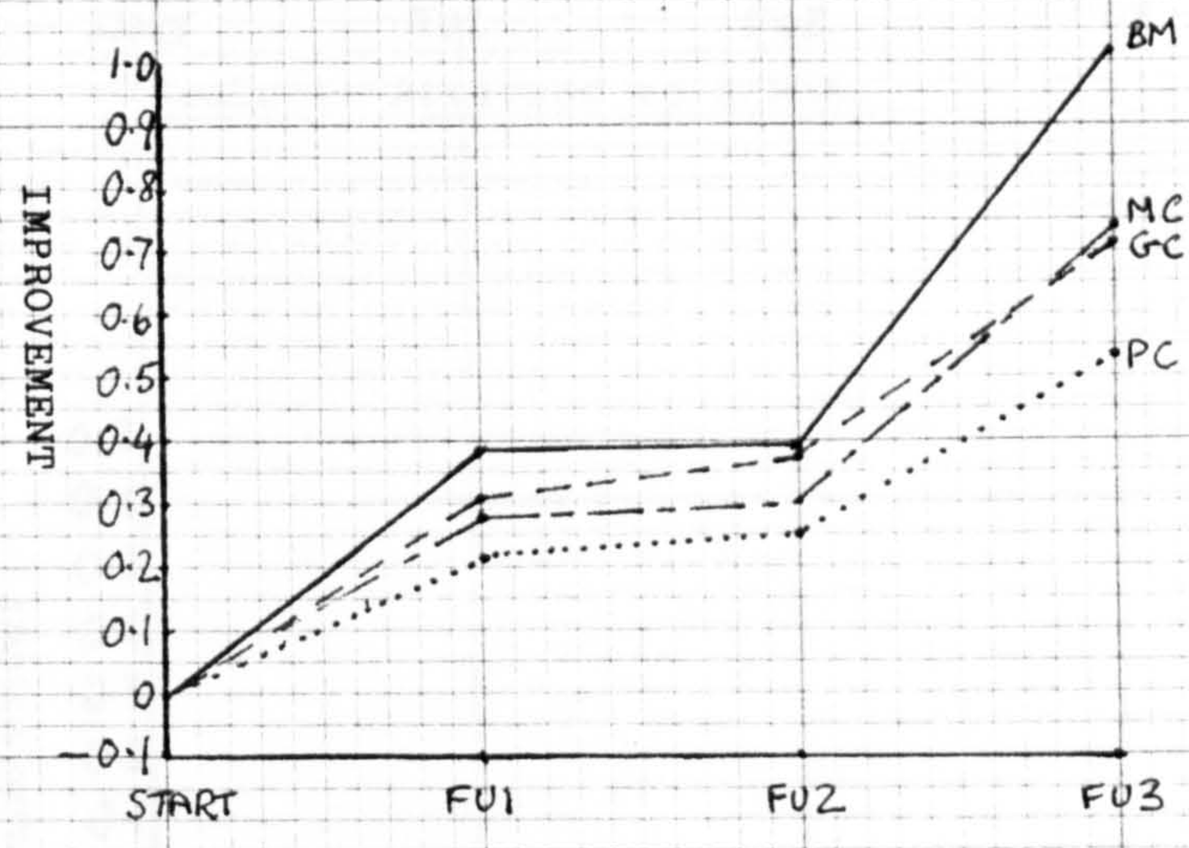


Verbal ability

Non-verbal activity

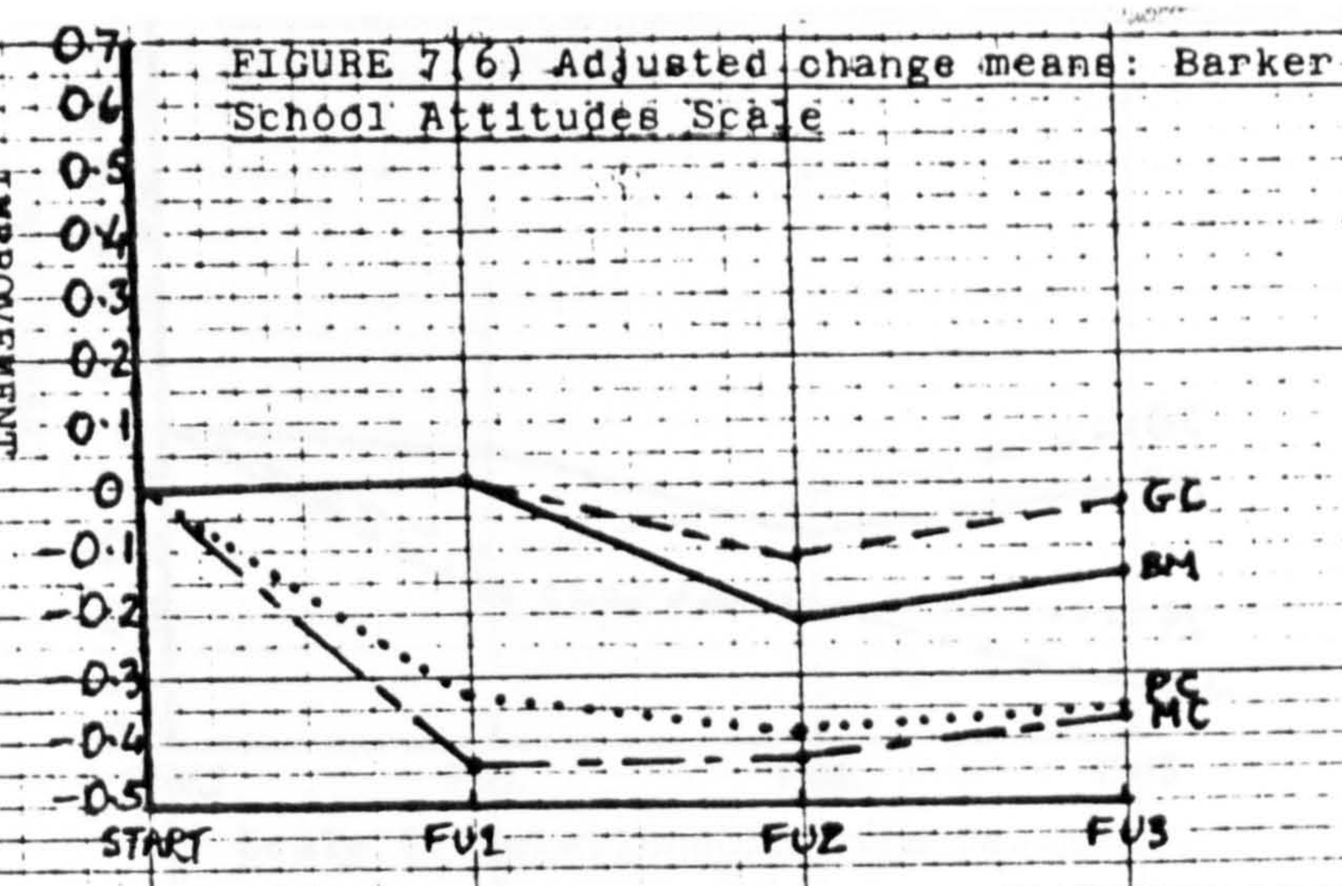
Total ability

FIGURE 7(5) Adjusted change means: reading comprehension

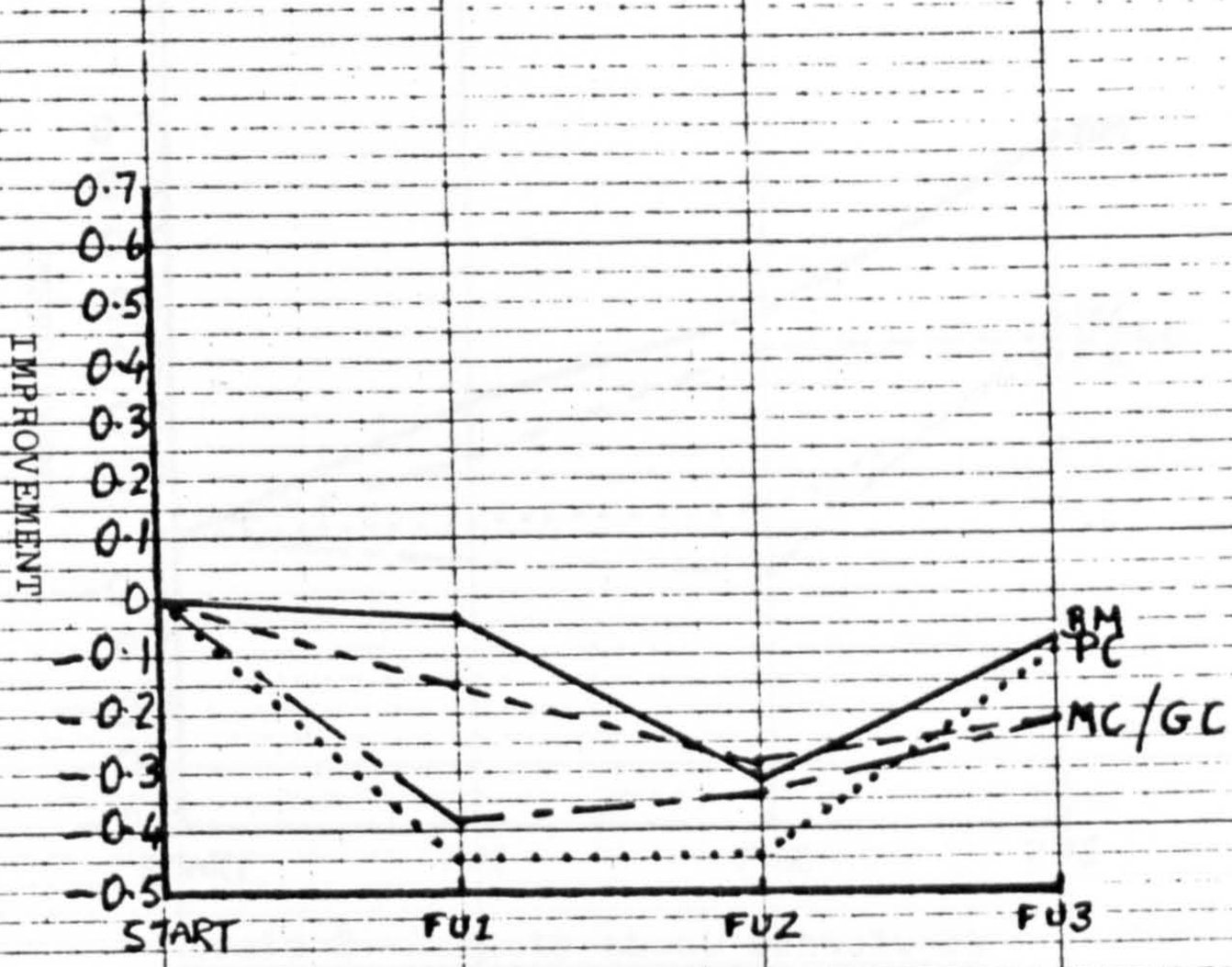


Reading Comprehension

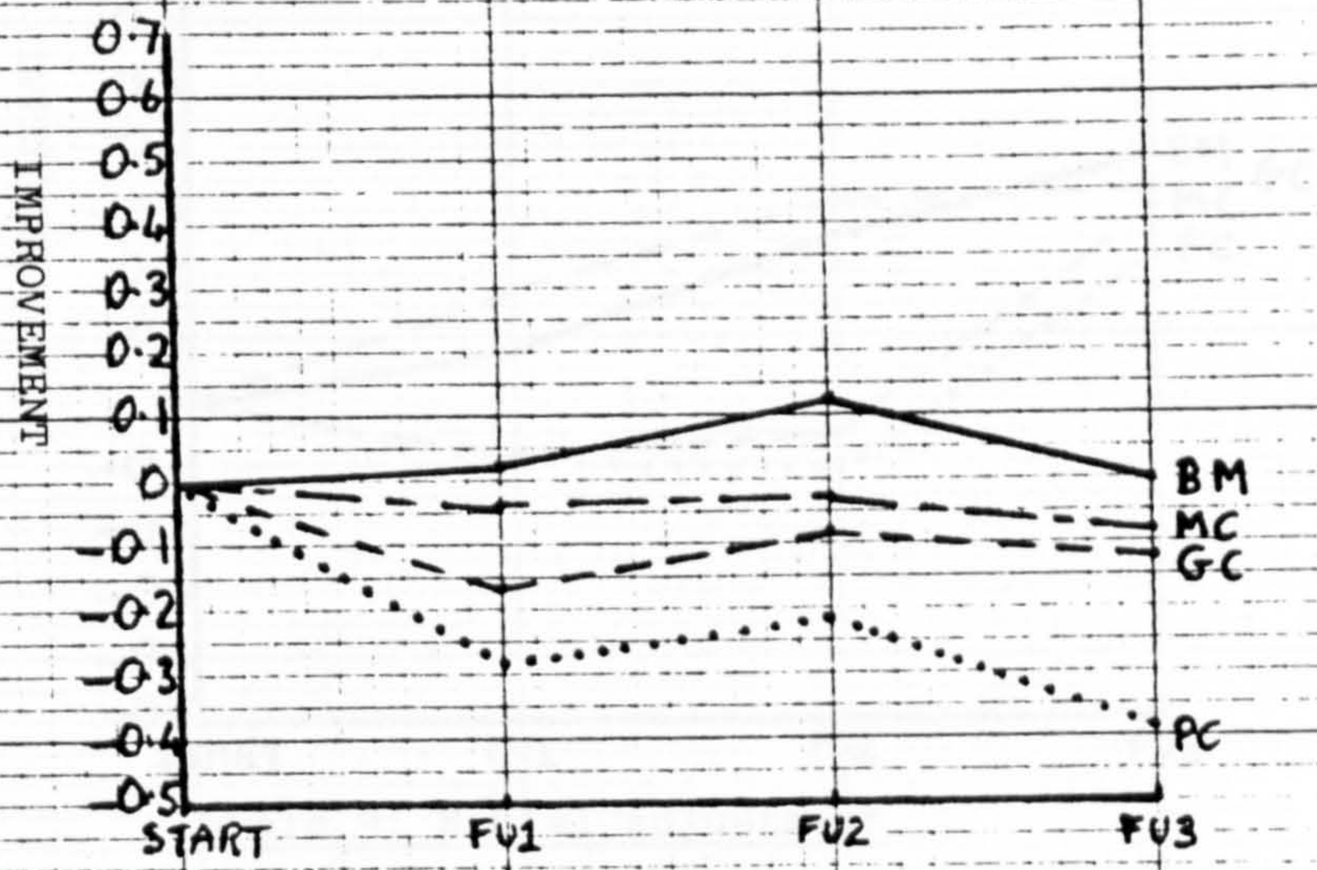
FIGURE 7 (6) Adjusted change means: Barker Lunn School Attitudes Scale



Scale 1: Attitude to school

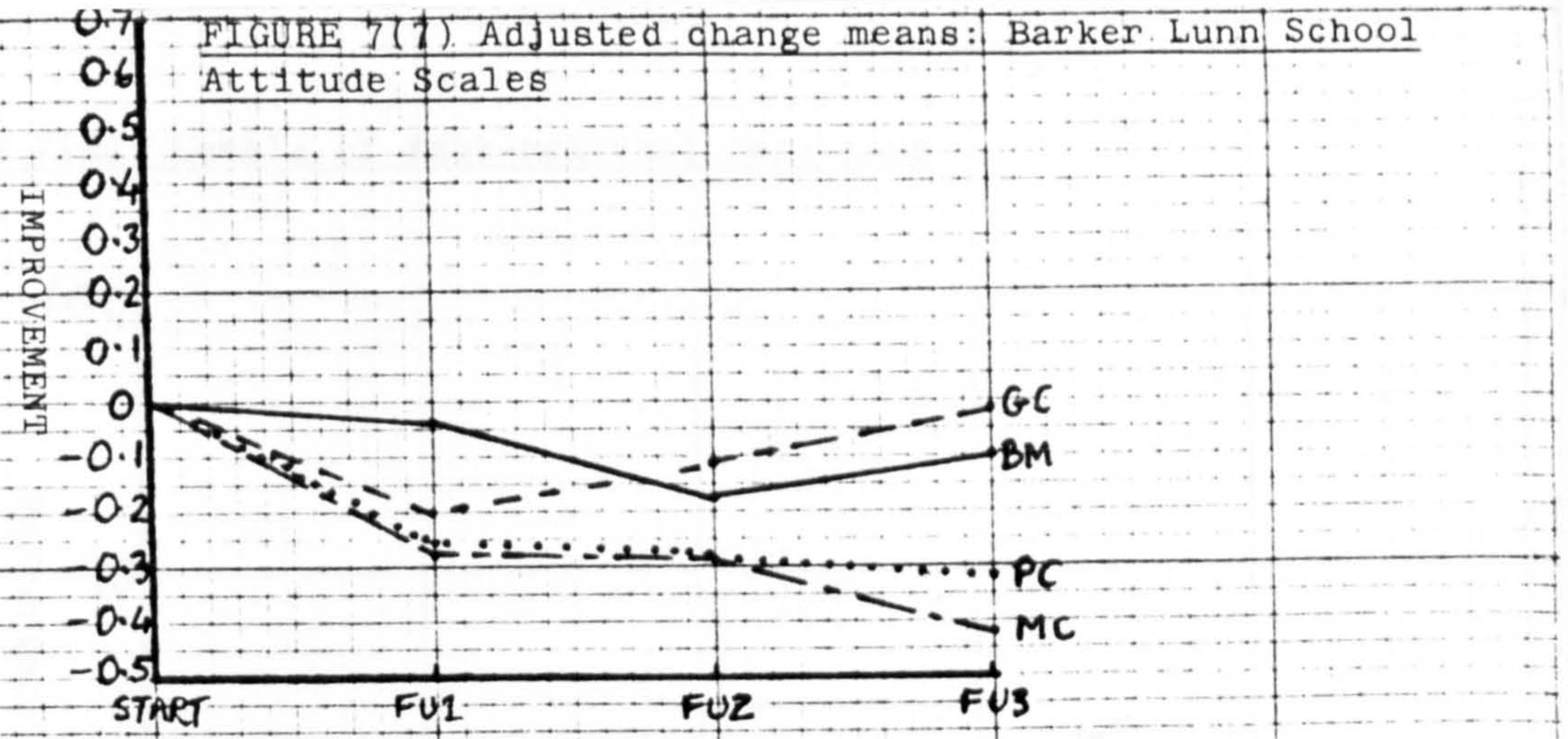


Scale 2: Interest in schoolwork

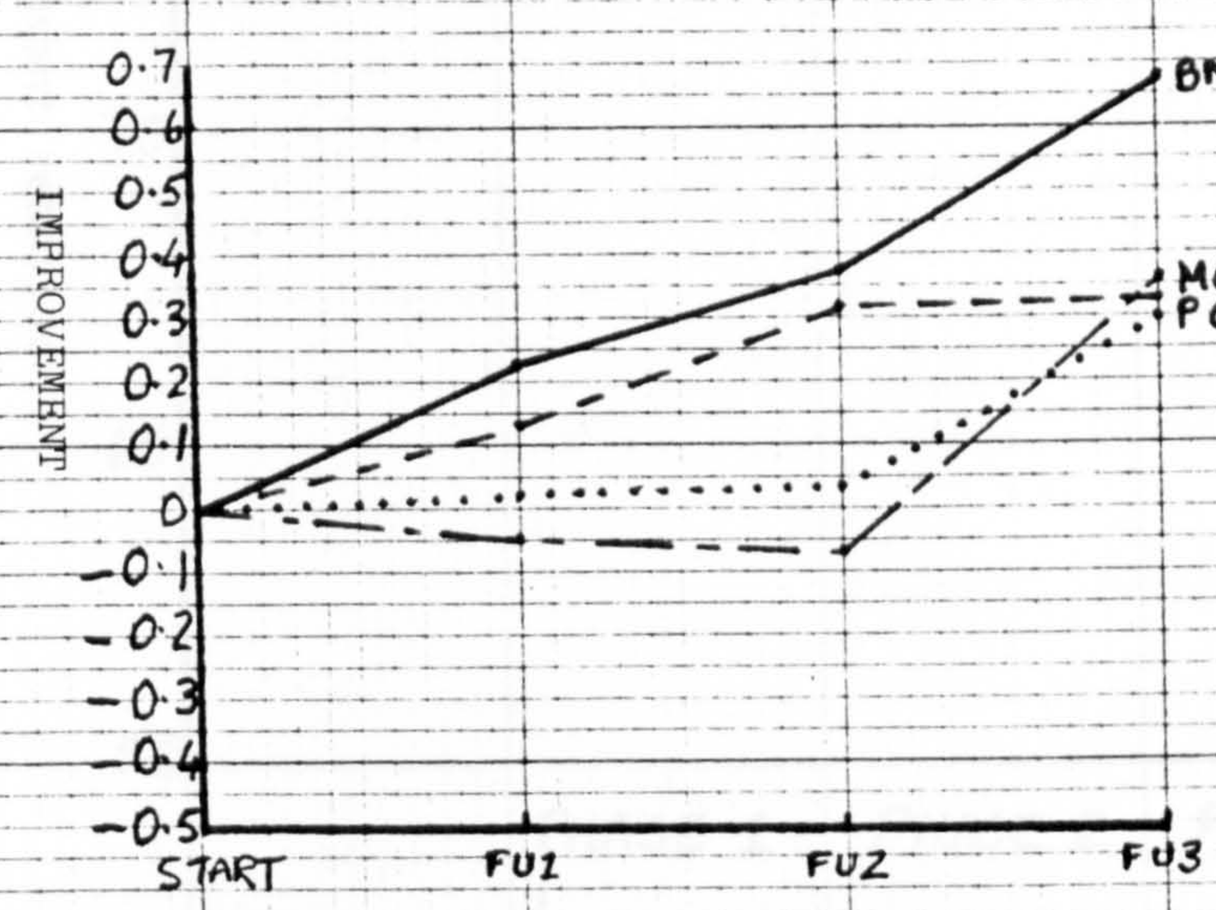


Scale 6: Conforming versus non-conforming

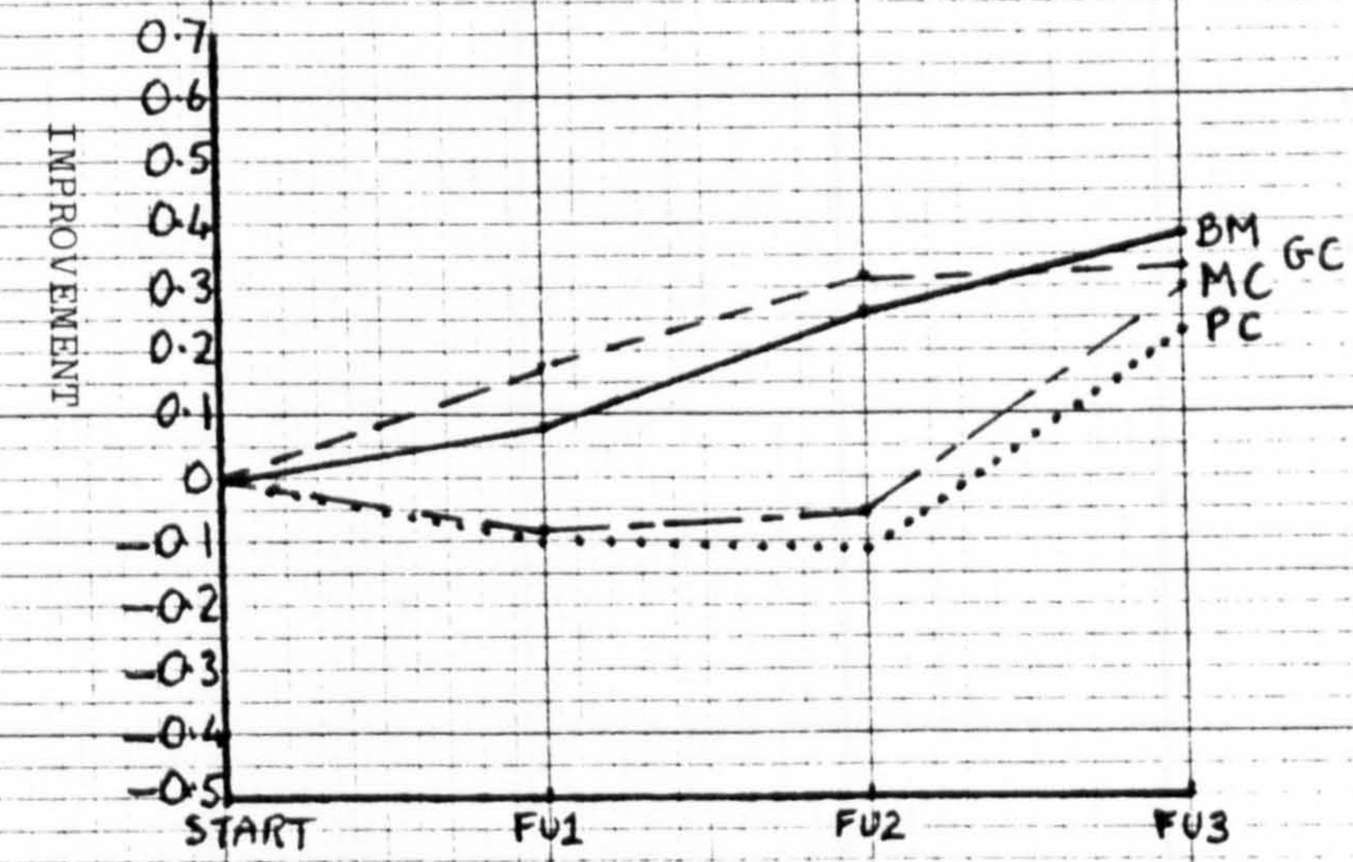
FIGURE 7(7) Adjusted change means: Barker Lunn School Attitude Scales



Scale 7: Relationship with teacher



Scale 8: Anxiety about schoolwork



Scale 9: Social adjustment

FIGURE 7(8): Levels of task-relevant behaviour

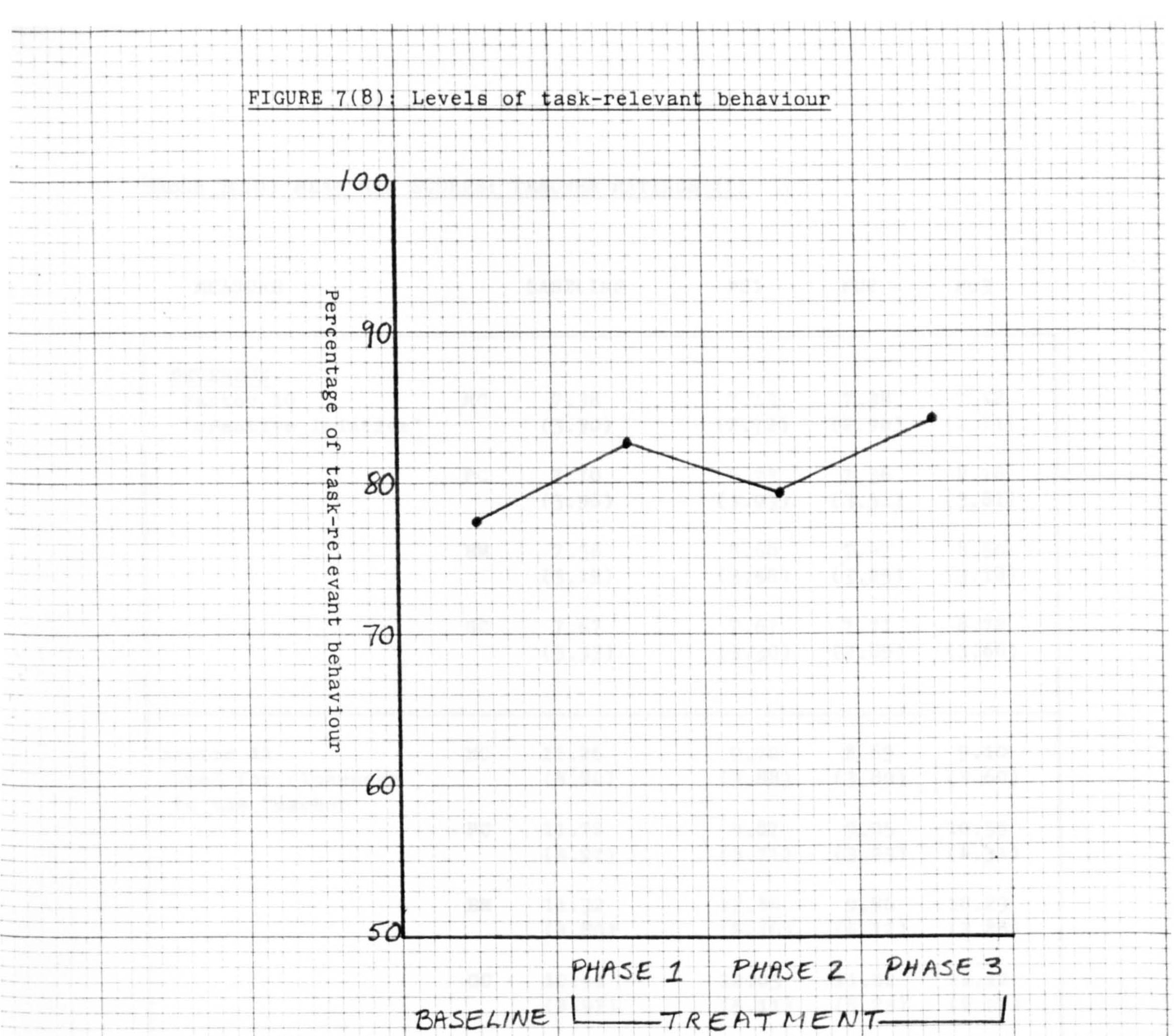


TABLE 7(10) MEAN RAW SCORES: TEACHER RATINGS (1)

MEASURE		BASELINE	FU1	FU2	FU3
DEVEREUX					
Factor 10 'creative iniative'	MC	7.38 (3.30)	7.13 (2.72)	7.38 (2.90)	7.48 (2.71)
	PC	7.70 (3.96)	6.73 (3.03)	7.36 (3.31)	8.18 (3.81)
	BM	7.36 (3.19)	8.60 (3.58)	7.47 (3.25)	8.56 (3.18)
	GC	7.67 (3.33)	7.67 (3.30)	7.71 (2.89)	8.72 (3.66)
Factor 11 'Need for closeness to the teacher'	MC	11.26 (4.90)	9.70 (3.88)	8.89 (3.80)	9.10 (3.60)
	PC	11.72 (4.97)	9.87 (3.71)	9.05 (3.23)	10.35 (4.31)
	BM	11.32 (4.80)	12.44 (5.16)	9.96 (4.17)	10.23 (3.88)
	GC	10.91 (4.97)	12.03 (5.11)	11.02 (5.23)	11.12 (4.29)

Note: 1) FU2 = first follow-up, FU2 = second follow-up, FU3 = third follow-up.

2) Figures in brackets are standard deviations.

3) Numbers of cases for MC, PC, BM and GC, respectively are: Baseline 92, 83, 74, 73; FU1: 75, 78, 70, 63; FU2 and FU3: 72, 69, 62, 54.

TABLE 7(11) MEAN RAW SCORES: TEACHER RATINGS (2)

MEASURE		BASELINE	FU2	FU3
RUTTER B2 Total score	MC	10.78 (7.66)	11.38 (10.17)	11.62 (9.38)
	PC	12.57 (8.12)	9.60 (7.68)	10.52 (9.89)
	BM	12.83 (6.86)	10.11 (8.26)	9.99 (8.18)
	GC	13.37 (8.3)	9.52 (7.91)	8.48 (7.39)
Neurotic score	MC	2.94 (3.25)	3.15 (3.15)	3.41 (2.74)
	PC	3.67 (3.21)	2.90 (2.94)	3.75 (3.63)
	BM	4.83 (3.50)	3.19 (3.25)	2.99 (2.47)
	GC	4.34 (3.45)	2.92 (3.15)	2.83 (2.40)
Anti-social score	MC	4.72 (4.52)	4.74 (5.22)	4.84 (4.91)
	PC	5.12 (4.53)	3.69 (4.05)	3.88 (4.36)
	BM	4.56 (3.61)	3.99 (3.79)	4.25 (4.31)
	GC	5.13 (4.06)	3.46 (3.51)	3.20 (3.84)

Note 1) FU2 = second follow-up, FU3 = third follow-up

2) Figures in brackets are standard deviations.

3) Rutter B2 data were not collected at the first follow-up.

4) Numbers of cases for MC, PC, BM and GC, respectively, are:

Baseline: 92, 83, 74, 73; FU2 and FU3: 72, 69, 62, 51

TABLE 7(12). MEAN RAW SCORES: SOCIOMETRIC AND PERSONALITY MEASURES

MEASURE		BASELINE	FU1	FU2	FU3
Isolation	MC	3.73 (3.06)	3.65 (2.84)	3.13 (2.98)	3.27 (2.77)
	PC	4.32 (3.44)	3.51 (3.28)	3.35 (2.77)	3.01 (3.09)
	BM	3.61 (3.33)	4.68 (3.51)	3.99 (3.14)	3.83 (3.01)
	GC	3.03 (3.12)	3.39 (2.80)	2.78 (2.42)	4.09 (3.09)
Neuroticism (JEPI)	MC	16.28 (5.45)	15.28 (5.34)	13.89 (5.82)	13.60 (6.31)
	PC	16.76 (5.33)	16.83 (4.87)	14.75 (5.83)	13.22 (5.90)
	BM	16.99 (4.99)	14.43 (5.37)	12.49 (5.10)	12.05 (5.66)
	GC	15.69 (5.56)	15.13 (5.74)	13.43 (5.58)	13.00 (5.54)

Note: 1) FU1 = first follow-up, FU2 = second follow-up, FU3 = third follow-up.

2) Figures in brackets are standard deviations.

3) Numbers of cases for MC, PC, BM and GC respectively are: Baseline: 92, 83, 74, 73; FU1: 81, 80, 68, 61; FU2 and FU3: 72, 72, 62, 51.

TABLE 7(13) MEAN RAW SCORES: ABILITY AND READING MEASURES

MEASURE		BASELINE	FU1	FU2	FU3
Verbal ability	MC	14.62 (8.18)	14.38 (6.95)	16.90 (8.72)	20.09 (11.89)
	PC	15.53 (8.87)	15.69 (8.84)	18.75 (10.19)	20.20 (10.00)
	BM	11.53 (6.10)	12.60 (6.73)	15.24 (7.99)	16.21 (9.43)
	GC	12.68 (5.41)	13.45 (5.55)	15.57 (8.24)	18.44 (10.01)
Non-verbal ability	MC	15.20 (6.68)	14.66 (5.03)	17.86 (6.58)	19.20 (7.57)
	PC	16.01 (7.33)	15.56 (7.41)	18.17 (7.76)	19.74 (7.72)
	BM	11.64 (5.67)	12.46 (6.52)	14.77 (7.67)	16.33 (7.90)
	GC	13.32 (5.42)	13.73 (5.47)	15.75 (7.20)	18.25 (8.22)
Total ability	MC	29.69 (13.60)	29.06 (10.89)	34.77 (14.06)	38.54 (16.07)
	PC	31.58 (15.50)	31.25 (15.62)	36.92 (17.07)	39.93 (16.41)
	BM	23.17 (10.87)	25.06 (12.39)	29.87 (14.73)	32.54 (16.77)
	GC	26.00 (9.89)	27.17 (9.97)	31.32 (14.81)	36.69 (17.57)
Reading comprehension	MC	13.24 (5.23)	14.75 (6.33)	15.40 (6.96)	17.37 (7.61)
	PC	13.19 (6.14)	14.59 (6.63)	16.26 (7.25)	16.62 (7.98)
	BM	10.65 (5.09)	12.91 (5.35)	12.90 (5.51)	15.84 (7.20)
	GC	12.44 (5.35)	13.66 (5.29)	14.22 (5.79)	16.07 (7.83)

Note: 1) FU1 = first follow-up, FU2 = second follow-up, FU3 = third follow-up.
2) Figures in brackets are standard deviations.
3) Numbers of cases for MC, PC, BM and GC respectively are:
Baseline: 92,83,74,73; FU1:81,80,68,61;FU2 & FU3: 72,72,62,51

TABLE 7(14) MEAN RAW SCORES: ATTITUDE MEASURES

MEASURE		BASELINE	FU1	FU2	FU3
Barker Lunn Scale 1	MC	2.90 (1.76)	2.01 (1.51)	1.99 (1.65)	2.11 (1.70)
'Attitude to school'	PC	2.30 (1.72)	1.94 (1.73)	1.92 (1.63)	2.08 (1.66)
	BM	2.54 (1.74)	2.50 (1.72)	2.10 (1.59)	2.28 (1.71)
	GC	2.63 (1.61)	2.76 (1.90)	2.45 (1.80)	2.54 (1.61)
Scale 2 'Interest in school work'	MC	2.71 (1.30)	2.11 (1.32)	2.20 (1.37)	2.22 (1.39)
	PC	2.46 (1.45)	1.83 (1.33)	2.13 (1.50)	2.45 (1.41)
	BM	2.18 (1.49)	2.39 (1.50)	1.96 (1.35)	2.24 (1.63)
	GC	2.60 (1.48)	2.42 (1.55)	2.20 (1.39)	2.31 (1.40)
Scale 6 'Conforming versus non-conforming'	MC	2.77 (1.36)	2.58 (1.10)	2.38 (1.23)	2.54 (1.13)
	PC	2.59 (1.42)	2.23 (1.22)	2.26 (1.27)	2.12 (1.24)
	BM	2.57 (1.41)	2.67 (1.25)	2.43 (1.30)	2.64 (1.10)
	GC	2.75 (1.41)	2.44 (1.31)	2.65 (1.36)	2.49 (1.29)
Scale 7 'Relationship with teacher'	MC	2.09 (1.47)	1.61 (1.41)	1.30 (1.28)	1.46 (1.26)
	PC	2.00 (1.47)	1.58 (1.22)	1.26 (1.11)	1.65 (1.23)
	BM	1.94 (1.27)	2.03 (1.37)	1.66 (1.53)	1.75 (1.53)
	GC	1.92 (1.47)	1.71 (1.34)	1.71 (1.44)	2.03 (1.44)

TABLE 7(14) MEAN RAW SCORES: ATTITUDE MEASURES (CONTINUED)

MEASURE		BASELINE	FU1	FU2	FU3
Scale 8 'Anxiety about school work'	MC	2.01 (1.48)	1.96 (1.52)	2.46 (1.61)	2.70 (1.60)
	PC	2.23 (1.78)	2.21 (1.56)	2.58 (1.59)	2.67 (1.68)
	BM	1.97 (1.29)	2.37 (1.60)	2.57 (1.51)	3.06 (1.50)
	GC	2.12 (1.63)	2.32 (1.74)	2.57 (1.51)	2.72 (1.70)
Scale 9 'Social adjustment'	MC	2.08 (1.46)	2.11 (1.30)	2.59 (1.27)	2.69 (1.27)
	PC	2.49 (1.34)	2.21 (1.37)	2.58 (1.32)	2.72 (1.22)
	BM	2.10 (1.19)	2.26 (1.18)	2.53 (1.20)	2.55 (1.23)
	GC	2.32 (1.51)	2.47 (1.35)	2.68 (1.39)	2.56 (1.50)

Note: 1) FU1 = first follow-up, FU2 = second follow-up, FU3 = third follow-up.

2) Figures in brackets are standard deviations.

3) Numbers of cases for MC, PC, BM and GC respectively are:

Baseline: 92, 83, 74, 73; FU1: 82, 77, 68, 60; FU2 and FU3: 79, 73, 65, 57

CHAPTER 8

DISCUSSION

DISCUSSION

Introduction

The present chapter has a number of objectives. First, interpretations will be offered for the findings reported in the previous chapter. This will include discussion of findings on specific measures, generalization, the relative performance of the different treatments, and the adequacy of treatment implementation. Second, some of the difficulties in introducing behaviour modification into ordinary schools will be examined, with reference to the complexity of the classroom and the demands on the teacher, teachers' acceptance of the behavioural approach, and the social context of such an intervention. Third, some critical views of behaviour modification will be presented. These will encompass both criticisms of behaviour modification itself as a treatment approach and also some wider objections relating to the model of scientific and research enquiry it espouses. Some suggestions will be made as to how the present kind of intervention could be modified to meet some of these criticisms. Finally, a number of methodological and practical recommendations from the study will be offered, based on interpretations of the findings and the wider issues indicated above.

(i) Interpretation of Findings

(a) General issues

The interpretation of the results of this study is not entirely straightforward for a number of reasons. First, there is the conflicting pattern of change observed. These patterns are inconsistent, whether viewed cross-sectionally, within assessment points, or longitudinally. Thus, within assessment points, there are discrepancies between change measures which might have been expected to show correspondence. However, the lack of agreement between measures is not entirely surprising, given the multifaceted nature of personality and behaviour and the likelihood that change in one domain may not permeate to another (Rosen & Proctor, 1981). Viewing results across.../

across follow-ups, gains are generally not maintained in subsequent assessments. These are the general trends in results and a closer examination of outcome on particular measures and groupings of measures will be conducted below which may contribute to a clearer understanding of patterns of change.

A second difficulty for interpretation has to do with the degree of success that is required to confirm the experimental hypotheses. At the level of supporting the hypotheses, success is minimal, as detailed in the Results section. However, as the hypotheses stand, extremely stringent demands are made of behaviour modification, in that it is, strictly, expected to out-perform the other two treatments. This expectation may be regarded as unrealistic and presumptuous and a more liberal criterion of 'success' could be adopted in terms of its performance relative to untreated controls, even if its advantages in this respect are shared with other treatments. Some modest claims can be made for it on such a criterion. In evaluating the impact of behaviour modification, it should also be borne in mind that an attempt was made in developing this regime to minimise research 'trappings' as much as possible, within the overriding constraint that it <u>was</u> a research study. This was done in an attempt to approximate regular school working conditions and typical psychologist-teacher consulting relationships, and to reduce those research factors which might contribute to positive outcomes (e.g. ample funds for reinforcers, special incentives for teachers to be involved, constant research staff support). The involvement of an independent observer was one conscious concession to research and measurement requirements. Most of the successes of behaviour modification are shared with group therapy, to a degree that is quite remarkable, as will be shown below, when the relative performance of the regimes will be considered. Both of these regimes, within the limits of the positive findings, consistently better the outcomes for parent counselling/teacher consultation.

(b) <u>Teacher ratings.</u>

Given that a major objective of the study was to increase children's task-relevant behaviour and diminish behaviour deemed inappropriate by the teacher.../

teacher, the limited impact on teachers' ratings is particularly disappointing. This is especially so since observational data suggested that the desired changes in classroom behaviour had to some extent been effected. Two Devereux dimensions ('Classroom Disturbance' and 'Inattentive-Withdrawn') seemed particularly relevant to assessment of these changes but significant differences emerged on neither, at any follow-up. On a more positive note, the significant differences in relation to all other groups on the Devereux factor 'Creative Iniative' and the shared advantage with GC over MC and PC on 'Need for Closeness to the Teacher' pointed to behaviour modification children having become more actively involved in classroom activities and valuing the teacher more. These improvements, however, were evident only at the end of treatment and not subsequently, and questions can also be raised as to the status of these positive outcomes as chance findings. If teacher approval for engagement in classroom activities can be seen as both increasing that engagement and leading to the child valuing that teacher more, then the disappearance of these gains following the end of treatment is perhaps not surprising, since the majority of the children were thereafter not exposed to the teachers involved in the study. In the absence of the approval to which they had been accustomed, behaviour they had shown relevant to these two dimensions would appear to have reverted to pre-treatment levels. No other significant differences were observed on the Devereux scale.

With teacher ratings on the Rutter scale, significant differences did not emerge until the final follow-up. Both advantages over MC (Total Score and Neuroticism) are shared with group therapy. It is tempting to attribute this gain to some accumulative process in which the benefits of treatment become evident only in the long term. However, if this was a genuine delayed effect of treatment, then it is not clear why teacher ratings on the Devereux scale showed no parallel changes. The lack of changes on the Devereux would at least appear to eliminate an interpretation in terms of 'demand' factors, with teachers being encouraged to report favourable change as a function of their involvement and investment in the programme or in keeping with what..../

what they felt the consultant desired. Any such effects would be expected to operate across both scales since teachers completed both scales. In any case, in view of the schools' practice of not having teachers continue with the same classes in the following school sessions, the teachers who completed behavioural ratings were not involved in the original programme so probably could not be regarded as systematically biased in their responses. An important consideration here is that the greater interval between treatment and follow-up, the more opportunity there is for uncontrolled influences to affect results. Thus, it is possible that differential therapeutic experiences in the intervening period may have favoured the behaviour modification and group therapy regimes. Controlling for such influences seems a well-nigh impossible task, given the range and diversity of experiences which may have a 'therapeutic' impact. Many of these, of course, may occur outside of conventionally-defined treatment frameworks, in the context of, for example, family, neighbourhood or community supports, friendships and so on. It is possible that these constitute confounding influences in many treatment studies.

(c) <u>Sociometric measures.</u>

Within the sphere of interpersonal relationships, a changing pattern of outcome was observed. On the Isolation measure, at the end of treatment, the behaviour modification regime showed significantly better improvement than all other groups. At midline, this improvement was maintained only in relation to GC. By the final follow-up, the superiority had disappeared and at this stage GC scores were significantly better than PC and MC.

Given the extent to which friendship choices depend on interactions and activities which take place outside the classroom, it is interesting for the behavioural approach that helping efforts which were essentially classroom-based did have this impact, albeit short-lived. Perhaps surprisingly, in view of the limited 'reach' of these efforts, somewhat greater change was observed on the 'like to play with' choice criterion, than on the 'like to sit beside' criterion. Changes on the more classroom-based criterion would have been more predictable.

The disappearance of the treatment advantage tends to suggest that gains made during the intervention phase were critically dependent on teacher support and reinforcement. Once these were withdrawn, the advantages over the other groups evaporated. These children may not therefore have learned the requisite skills for making and sustaining friendships which could underpin their interpersonal functioning beyond the time when adult support and encouragement were forthcoming. If this interpretation is correct, then the present pattern of results points to a possible limitation of helping methods which hinge on adult reinforcement, as pointed out by Combs & Slaby (1977).

(d) Personality ratings.

Some of the concerns surrounding the place of personality measures in an assessment battery in a behaviourally-orientated model were expressed earlier in the section dealing with screening methods. One important reservation concerned the possible clash between assumptions of enduring states or traits in personality measurement on the one hand, and assumptions of the situation-specificity of behaviour in behavioural models on the other. The use of a personality measure in a behavioural study, such as the present, does not, however, imply an acceptance of the former set of assumptions. While interpretations of the questionnaire responses may differ from those intended by its author, they are not thereby invalidated. The pattern of changes observed on JEPI neuroticism suggests that this kind of measure is of some relevance to an evaluation of behaviour modification. Thus, BM shows significantly lower levels of neuroticism than PC at the end of treatment, and at midline, along with GC, it shows lower levels than both MC and PC.

Given the generality of reference of the items in the JEPI, the observed changes can be taken to reflect improvement in children's perceptions of their wider functioning beyond the classroom. Such an interpretation would be consistent with a report by Best (1973) who found that even when pupils completed the JEPI in the classroom, they answered 80 per cent of the items with reference to their feelings at home. Children spend a considerable portion of their waking day in and around school so it is perhaps not surprising that.../

that more generalised self-reports should show beneficial changes if there is greater satisfaction with classroom life, and relationships with teachers and peers are improved - as changes on the Barker-Lunn, Devereux and sociometry would tend to suggest.

(e) Attitude measures

It is of interest that, at the midline assessment, at the point when neuroticism was reduced in the behaviour modification group (along with GC) relative to MC and PC, both these groups also showed a reduction in the attitude 'Anxiety in class' relative to MC. This improvement in expressed attitude suggests a linkage between feelings of anxiety specific to the school setting and anxieties experienced more widely. However, on neither measure did the improvements persist till the final follow-up.

Lack of continuity in change in the present results is perhaps most apparent on inspection of the Barker Lunn scales across assessments. Although five different scales are implicated in significant comparisons between groups - with behaviour modification involved in all five - none of the scales throws up significant results at a subsequent assessment. An indication of the inconsistency between measures from different sources is that despite the improvement in sociometric status by the end of treatment, no parallel change was recorded in the attitude dimension 'social adjustment - getting on well with classmates'.

Similarly, although one Devereux dimension suggested an improved relationship with the teacher, at the end of treatment, the attitude dimension 'Relationship with teacher' showed no corresponding change. In both these examples of inconsistency between measures, the data were derived from different sources. Even within the attitude scale itself, however, some expected associations in change did not materialise. Thus, although 'Attitude in school' improved, the related dimensions of 'Interest in school work' and 'Importance of doing well' did not, or at least, in the case of the former, not to the same degree (the advantage was only in relation to PC). In the intercorrelations reported by Barker-Lunn (1969), these scales are highly related..../

related. Discontinuities apart, there is evidence that the present behavioural intervention did have an impact on children's attitudes, improving attitudes to school and reducing perceived anxiety in class relative to MC, and improving interest in school work and perceptions of social adjustment and conforming behaviour relative to PC.

(f) Ability measures

The behaviour modification and group counselling regimes both showed improved verbal, non-verbal and total ability scores at the midline assessment, relative to both MC and PC. No significant gains were apparent before or after this stage. The content of these measures cannot be regarded as having obvious counterparts in routine classroom activities which could be influenced and improved with direct implications for test performance. How therefore may these be explained?

One possible interpretation is that they could be regarded as reflecting a general improvement in functioning with performance on such tests being one index of this. Support for this interpretation comes from consideration of other changes observed at the same time, namely reduction in JEPI neuroticism and 'anxiety in class' on the Barker Lunn. The favourable changes observed at end of treatment may also have been contributory. These improvements would therefore be regarded as in some way facilitative of test performance.

Another, possibly complementary interpretation, might implicate improved motivation for such exercises and the earlier improvement in 'attitude to school' could be seen as supporting this. There is no evidence that general improvement in functioning - if that indeed occurred - was any greater at the second follow-up than the first. So, other than invoking some cumulative process of improvement, it is not clear why the significant differences emerged when they did.

(g) Reading comprehension

On measures of reading comprehension, the performance of the behaviour modification group proved disappointing. At no point was superiority over untreated controls demonstrated, and the only advantage observed was in relation..../

relation to PC at the final follow-up. Furthermore, no relationship was established between an increase in task attention and improvement in reading comprehension scores. This negative outcome is in keeping with Lahey's (1977) criticism that where reading comprehension is taken as the index of reading achievement, there is little evidence that achievement can be influenced by positive reinforcement. The failure to effect improvements is also consistent with those studies which have shown that a focus on increasing task attention may not be sufficient to induce changes in academic performance (e.g. Ferritor et al., 1972; Harris & Sherman, 1974). In the present study task attention was the primary but not the exclusive focus where study behaviour was concerned. Another consideration here is the 'strength' of the reinforcement employed. It is possible that token reinforcement with attractive back-up rewards may have been more effective. As indicated in the literature review, few studies have relied solely on social reinforcement when academic performance has been the target.

(h) Classroom observations

The most direct measure of the impact of the intervention procedures is that provided by classroom observations. The results indicate that task-relevant attention was significantly better than the baseline in the first and third phases of the intervention, but not in the second. How much confidence may be placed in these findings? While the observational procedure itself may be regarded as reasonably reliable, there are other sources of possible error. First, in the absence of data on control subjects, the influence of non-experimental variables cannot be ruled out. It is quite possible, for example, that at different stages in the term there might be variations in the level and quality of pupil attention and application to work. For example, task attention could well be better in the later stages of a term or session than in the earlier stages when children are accommodating to classroom rules and some testing-out might occur.

A..../

A second difficulty in interpreting the observational data is that the observer could have unwittingly biased the data, since she was aware of the aims of the intervention. Although she was not involved in any discussion of data trends during the intervention phase, knowledge of the aims of the programme could not be prevented since she was conversant with the study design. Observer expectancy effects could not therefore be ruled out.

In the light of these cautions, the observational data can be regarded as only suggestive rather than definitive. If, however, they are taken at face value as indicating a 'real' treatment effect, what do the changes mean?

In the absence of norms for desirable levels of 'task-relevant' behaviour, or evidence in the present study that rises in such behaviour are associated with achievement gains, the significance of the changes observed remains unclear. In a study such as that of Becker et al. (1967) where task relevant behaviour rose from a baseline level of 38% to 71%, the change seems dramatic and obvious. In the present study, the baseline level of 77.5% rose to 82.7%, 79.5% and 84.1% in the three treatment phases. It is unlikely that, as a group average, these improvements are of much clinical as opposed to statistical significance. It is also possible that changes of this order would not be readily detected by teachers, which may help to explain the relative lack of changes on the behavioural ratings they provided. With regard to the degree of impact of the programme, it may be that the baseline level of task-relevant behaviour - in terms of putative norms - was already quite high so that a ceiling effect may have limited the room for change. It is also to be noted that the pre-treatment value may not have constituted a 'true' baseline, since these observations were contemporaneous with the first phase of training, and some experimentation with procedures may have begun.

How may the rise, fall and recovery of the group average for task-relevant behaviour across treatment phases be explained? One possibility is that the teachers, or the consultant, or both, may have slackened off their.../

their efforts after the early phase of treatment, but restored the previous level of commitment in the last weeks of the programme. It is hardly surprising that where treatment requirements demand considerable effort and organisation by teachers, in the face of many other competing pressures, their application should prove to be variable. Other investigators (e.g. Hersen, 1981) have commented on this. Initial high praise rates, for example, may not be sustained and reverting to inappropriate disapproval may occur. However, there were no indications from either formal or informal observation to suggest that the nature of teacher attention - in terms of enthusiasm, quality or quantity - varied systematically across the three treatment phases. From the children's point of view, it is possible that some degree of habituation to teacher attention could account for the fall in level of task attention in the middle phase.

With regard to teacher behaviour, no significant changes were detected in rates of approval and disapproval across baseline and intervention phases (see Appendix 7). It is therefore not possible to relate the observed changes in child behaviour to variations in the nature of teacher attention. As pointed out in Appendix 7, a wider sampling of teacher behaviour rather than simply in relation to target children, may have detected some changes.

In view of the restriction of observation to the baseline and intervention phases, it is not known how well the final improvement was maintained. Given that the end of treatment occurred at or near the end of the school session, and that the children transferred teachers in the new session, maintenance was unlikely.

Given that the thrust of the present study was towards intervening in the 'natural' setting of the ordinary school environment with as little as possible in the way of research trappings, the placing of an independent observer in the classroom was a concession towards research and measurement requirements. It was accepted that such personnel are rarely available in ordinary schools. Teachers who are concerned with measurement and evaluation have to evolve recording procedures which are compatible with ongoing teaching activities.../

activities, or else encourage children to self record (McNamara, 1979).

Having considered the findings for individual measures, the next two sections will attempt an overall view of the results of the study by examining (a) the findings for behaviour modification from the perspective of generalisation and (b) the relative performance of the treatment groups.

(j) Generalisation

It will be recalled that a distinction can be made between three types of generalisation: temporal generality, or maintenance of treatment effects after the termination of treatment; setting generality, or the occurrence of treatment effects in settings other than the ones in which treatment has been conducted, and behavioural generality, or changes in behaviours not the specific target of treatment interventions.

As the foregoing discussion of results will have demonstrated, there is little evidence of temporal generality in relation to the behavioural intervention. Taking a stringent interpretation of maintenance of effects over time would call for superiority to be shown on the same measure across follow-up. None of the comparisons in relation to the untreated controls showed this. The only evidence of such maintenance was in relation to PC in the case of neuroticism and in relation to GC in the case of isolation, both occurring from first to second follow-up. A less stringent criterion for maintenance might view evidence of positive outcome at later follow-ups, where none has been observed before, as constituting a retention of treatment benefits which reach significance only in the long term. With this more lenient criterion,'maintenance' is demonstrated (in relation to MC on neuroticism, ability measures and 'anxiety in class' at second follow-up, and on teacher Rutter ratings at third follow-up). The reduction in number of positive comparisons for BM across time would seem to point to diminishing treatment effects the later the follow-up. Confidence in the more generous interpretation of maintenance has to be tempered, of course, by consideration of how factors outside of experimental manipulations may affect outcome in the longer term, as discussed earlier.

Even if the less demanding interpretation of maintenance is accepted, perhaps little more than the effects observed could be expected, given the factors militating against carry-over of treatment effects. The long summer holiday following the intervention and the transfer to new teachers in the new session may have helped to dissipate the benefits of treatment. In addition, although written guidelines were made available to teachers in the new term, it is unlikely that such material alone would be sufficient to affect teacher behaviour.

The observation of benefits in the longer term, at the final follow-up, are of some interest in the light of the findings of Wright et al. (1976). It was noted in the earlier review of child psychotherapy studies that these writers had found better outcome at follow-up than at treatment termination in their re-examination of controlled studies as had Levitt (1957). In the context of treatment by psychotherapy, such findings are consistent with the notion that psychotherapy achieves its results by changing 'core' features of personality and that these improvements are gradual and not necessarily evident in the short term. Wright et al. (1976) advance such an argument to explain their own findings. Theoretically, treatments with a behavioural orientation can also have long term effects if one views the intervention as significantly changing the child's environment or improving the child's skills in coping with it (Rutter, 1982). Such changes might have cumulative effects, for example, as a child's improved functioning elicits approval, which, in its turn, motivates the child to further improved performance. For whatever reasons the positive benefits of the present study were apparent at the final follow-up, they should provide encouragement to behavioural researchers to monitor the outcome of their interventions in the longer term.

Data are not reported in the present thesis which could have a direct bearing on the question of setting generality. This issue is, however, inextricably bound up with the question of temporal generality. Since children changed teachers and classrooms after the end of treatment, evidence for 'maintenance'.../

'maintenance' in the longer term is also partly evidence of setting generalisation. The distinction made here between the three different types of generalisation is, of course, merely a convenient theoretical one rather than a reflection of any independent reality.

Where behavioural generality is concerned, the earlier discussion will have indicated that a number of changes were demonstrated in processes which were not the direct targets of intervention. These occurred in, for example, ability and personality measures, and in attitude components. It is intriguing that positive changes were observed within all the domains of measurement adopted at some stage in the follow-up. Although the number of significant changes so recorded was rather limited, the 'spread' of changes across different measures provides some support for the adoption of multiple measures and perspectives on the process of change, rather than reliance on one, as urged by a number of writers (e.g. Johnson & Eyberg, 1975; McNamara, 1975).

To demonstrate correlated changes between behaviours is one thing: it is a different matter to demonstrate how or why behaviours are related in such a fashion that change in one has implications for the other. It has recently been stressed that the field of therapy evaluation has a great need to develop a conceptual framework which can approach the issue of correlated changes systematically, rather than merely record their occurrence. Thus, Rosen & Proctor (1981) distinguish between different kinds of therapeutic outcome and identify 'instrumental' outcomes as ones where certain kinds of change are preconditions for the 'ultimate' or terminal goals, or in some way facilitate them. For example, in the present study the hypothesis was advanced - but not substantiated - that an increase in task attention (instrumental outcome) would have implications for reading achievement (ultimate outcome). Rosen & Proctor's argument that all outcomes employed in the evaluation of treatment should be tied to ultimate outcomes through an explicit rationale is but another facet of the argument for a more systematic approach to the conceptualisation of correlated changes in behaviour. In a similar vein, Kazdin.../

Kazdin (1982) writing in the context of symptom substitution, has urged systematic investigation of response covariation. He suggests that examination of individuals' structure of experience in relation to behaviour and situations can enable predictions to be made about specific relationships in behaviour change. Such examinations and predictions of response relationships have not been undertaken in the present study. However, some of the interpretations which have been offered of possibly related changes (e.g. neuroticism and attitude changes contributing to improved ability score) might suggest a basis for further work to explore such linkages.

The generalisation issues considered above may be regarded as a subset of wider generalisation questions which may be addressed to the study as a whole. In this wider perspective, one is concerned with the 'external validity' of the study, that is, the extent to which its findings can be generalised outside its own specific context, for example, to other populations, settings, behaviours and treatments (Campbell & Stanley, 1966). In general, having intervened in natural or regular school settings can be seen as assisting the external validity of the study, though not guaranteeing it, in that it is to these settings that one would wish the findings to be generalised. Some comments can be offered on the different components in generalistion. With regard to generalisation to other populations, firstly, it was felt that the identified group of children comprised a fairly representative sample of children experiencing problems within ordinary schools. Clinical ratings also suggested that the selected children were experiencing genuine problems. While these considerations, and the heterogeneous nature of the presenting problems, may be regarded as assisting generalisability, they must be balanced, on the negative side, by consideration of withdrawals of cases both before treatment commenced and during the process of treatment and follow-up. Outcome data may have been analysed for a biased sample, which would prejudice generalisability. While comparisons of available data for those who dropped out and those who remained showed no differences, this does not preclude the possibility of there being other important differences, not tapped by the.../

the available measures, which could have been related to outcome. With regard to the population of teachers, rather than children, some confidence can be placed in external validity. With teachers being selected by the head teacher on the basis of likely motivation and their exposure to target children, a wide range in attitude and motivation was observed in the group. The teachers' group could not be regarded as uniformly enthusiastic and ardent proponents of the behavioural approach. In this respect, the teachers were probably fairly typical, and their range of response is likely to mirror the regular school situation, where teachers can rarely be handpicked for their involvement.

With regard to generalisation to other behaviours, changes were observed on a spread of measures, reflecting some degree of impact in different areas of functioning rather than in a restricted sphere. This breadth of effect may therefore offer confidence as to wider generalisation.

Finally, does the treatment itself have external validity? There are two important considerations here. First, while there were many teachers applying the treatment procedures there was only one consultant. In view of the apparent importance of consultants' support and encouragement of teachers (e.g. Brown et al., 1969; Rollins & Thompson, 1978) and the ultimate implications of this for children's behaviour, experimentation with a number of consultants would have been preferable. This would have tested the dependence of outcome on the conduct of treatment by the single consultant. Within the practical constraints of the present study, this more elaborate design was not possible. A second issue bearing on the replicability of the treatment is that in a 'package' form, rather than in a form emphasising a single procedure, it is difficult to identify the effective component(s). An attempted replication may fail to include the components which were actually responsible for the effects. Azrin (1977) argues that such 'package' approaches should be adopted unapologetically in order to obtain, ideally, a total treatment success. Once sucess is clearly established, then analytic studies of the programe can be undertaken. Such an approach is probably justifiable where one is striving,.../

striving, as in the present study, for therapeutic effects with multidimensional problems. However, it does not make for a straightforward translation in an intended replication. This is perhaps one of the 'costs' for external validity in an applied study in a natural setting.

To conclude this section, it should be pointed out that the above comments on external validity are offered in a speculative fashion, since generalisability cannot be taken for granted but has to be tested for empirically.

(k) The relative performance of the treatment groups

In examining the respective performance of the three treatment groups, a number of noteworthy features call for comment. First, it is apparent that, on the measures of effectiveness employed here, and in the form in which it was applied, parent-counselling/teacher consultation proved ineffective. Where significant differences emerged between groups it was consistently bettered by behaviour modification and group counselling. In one case, it did worse than untreated controls (JEPI neuroticism at midline). The description of teacher response to this programme (see Kolvin et al., 1981) with 62% of the teachers involved making a questionnaire return, suggested that over half the teachers considered consultation to be of at least some use in improving their pupil handling ability, in increasing their knowledge of psychological methods of managing children, and in thinking out alternative ways of coping with disturbed behaviour. Over three-quarters of the teachers reported finding the social workers' support helpful. Although the return rate for the questionnaire was low, and questions the universality of these findings, it leaves the impression of a large gap between 'success' estimates at consultee level and at client (pupil) level. In this regard, Cowen (1978) has argued that change criteria at consultee level, while important, are not sufficient. Even at the level of consultee data, Schmuck (1968) found no changes in behaviour despite significant cognitive and attitudinal changes.

What factors might underlie the difference in effectiveness between behaviour modification and parent counselling/teacher consultation, despite their sharing of a teacher consultation component? Although data relating to treatment

implementation...../

implementation do not allow clear differentiation of the two models, the writer would propose the following emphases as giving behaviour modification the advantage in the present comparison: 1) specifying clear and discernible features of the child's behaviour on which the teacher could focus; 2) developing successive objectives in a 'shaping' process; 3) stressing contingencies between changeworthy behaviour and positive consequences; 4) the stress on <u>increasing</u> teachers' rates of positive responses and encouragement. The apparent superiority of the behavioural approach in consultation in the present study is consistent with the outcome of similar comparative studies (Jason & Ferone, 1978; Jason et al., 1979) and also with the conclusions of reviews of varying approaches to teacher consultation (e.g. Medway, 1979).

A second important feature of relative group performance concerns the similar impact of behaviour modification and group counselling. In terms of a crude tally of successes, behaviour modification is superior to untreated controls in the present study on eleven measures, group counselling on twelve. The respective performance in relation to parent counselling/teacher consultation is twelve successes for behaviour modification, eight for group counselling. What is much more intriguing is that of the fifteen separate measures in which one or other records an advantage over the controls, eight are shared by them. Furthermore, where change is shared, it occurs <u>only</u> at that shared point, and at no other follow-up point for either of the regimes independently. This coinciding of improvement covers four measurement domains: teacher ratings (both Devereux and Rutter), personality, attitude and ability measures. This similarity in impact is remarkable given the theoretical divergence between the two models. It is less surprising, however, in the light of three considerations. First, the measures were selected to stand free of any particular theoretical approach rather than being tied in favour of particular ones. Second, regardless of differing theoretical and treatment orientations, the children as treatment clients are centrally involved in making <u>their</u> own sense of their experiences, and integrating the various influences to which they are exposed. Children within different..../

different regimes may therefore show convergences in their perceptions, attitudes and behaviours which belie the discrepancies between models. Third, and relatedly, the children in the two regimes may have shared the experience of having a significant adult displaying interest, attention and concern, although this may have been demonstrated in rather different ways. In behaviour modification, considerable emphasis has been placed on techniques or procedures as the major determinant of outcome, while in traditional psychotherapy the therapeutic relationship has been given a more central role. These distinctionns seems to be less clear now than they used to be and the role of relationships in behavioural work has recently been given more attention (DeVoge & Beck, 1978). With regard to the present discussion, it is of interest that Sloane et al. (1975) showed that in some instances behaviour therapists scored higher on Rogerian interpersonal variables than did supposedly traditional psychotherapists. Additionally, analysis of non-directive counselling has revealed much in the way of reinforcement processes (Truax, 1968). These considerations would suggest that in view of the commonalities in outcome of different treatments, research into children's experiences and perceptions of therapy and therapists would be of considerable value. Very few comparative treatment studies have, in fact, been conducted with children. It is of interest that in two other studies, comparing behaviour therapy and psychotherapy (Humphrey, 1966; Miller et al., 1972), albeit with different populations from the present one, marked divergences between treatments were not found.

In considering the parallel successes of behaviour modification and group counselling, the consistent differences between them and untreated controls would seem to belie Levitt's (1957) pessimistic conclusion that child therapy offers little that can better improvement rates without treatment. In addition, it is evident that both these approaches have the potential of reaching large numbers of children. As suggested in earlier discussion of the numbers of children experiencing difficulties within the ordinary school framework, the question of 'reach' becomes a very relevant concern in the planning of helping interventions..../

interventions.

In considering the pattern of results in general and the apparently greater degree of success of behaviour modification and group counselling, relative to parent counselling/teacher consultation, it is important to pay attention to possible biases which may have affected the outcome. A number of issues need to be addressed here. First, could biases have existed in children's responses or teachers' ratings? With the former, self-conscious involvement in a 'treatment' process was perhaps most characteristic of children in group counselling, so responding in the expected direction, for example, on personality or attitude measures, might be most predictable in this regime. However, as many positive outcomes are recorded on such measures for behaviour modification as for group counselling, so bias seems unlikely. With teacher ratings, the mere fact of direct involvement in a regime, as in behaviour modification and parent counselling/teacher consultation, does not appear to have produced biased responding. Positive results appeared only with the former regime, and not, indeed, until a follow-up point when the 'trained' teachers were no longer involved with the target children.

A second consideration is that the two apparently more successful regimes were conducted in the second year of the research, and the unsuccessful one in the first. The initial levels of disturbance did not differ between the two years (Kolvin et al., 1981) so this factor could not have contributed to the divergent outcomes observed. Another possibility is that by the second year of the project the school environments had become more 'therapeutic' so that additional non-specific help, unrelated to their treatment regime, was available to second year target children. This possibility, however, seems extremely unlikely. If indeed the school environment benefitted in this fashion, then it is likely that children treated in the first year would have shown some advantages by the final follow-up, but they did not do so.

A third potential bias in the pattern of outcomes concerns the issue of attrition of cases. Where differential attrition occurs across regimes, it.../

it is possible that advantages accrue to particular treatments, if, for example, the sample remaining is composed of the more motivated and less resistant cases. It was noted in the Results Chapter that the highest attrition rate during treatment follow-up was in the group counselling regime. Comparison of baseline data for dropouts with data for those continuing in treatment did not show significant differences. The possibility remains, as indicated previously, however, that differences related to outcome may have existed at other levels not tapped by the available range of baseline measures, so the possibility of attrition bias in favour of group counselling cannot be excluded.

The discussion so far has stressed the similarities in effectiveness of behaviour modification and group counselling. What about the differences between them? The 'successes' which the two treatments do not share tend to make for a difference in the phase of follow-up when effectiveness is most apparent. Thus, behaviour modification has most successes in the period covered by the first and second follow-ups, while group counselling appears most effective across the second and third follow-ups. A somewhat similar trend was observed, within a much shorter time span, by Marlowe et al. (1978). They found that reinforcement procedures were associated with early gains in task attention, and counselling procedures with improvements towards the end of the intervention period. Another important difference is that in no comparison does group therapy emerge with a statistically significant superiority over behaviour modification. However, behaviour modification betters group therapy on improvements in 'creative initiative' and isolation at the end of treatment, and isolation again at the next follow-up. Behaviour modification's early superiority with regard to isolation nicely illustrates the earlier point about particular phases of effectiveness because, at the final follow-up, it is group counselling rather than behaviour modification which demonstrates superiority over MC and PC. This resurgence, however, did not achieve statistical significance relative to behaviour modification. Despite some differences in time of best impact, there is no compelling evidence that either treatment was specifically effective with particular domains of functioning in terms of inducing changes which persist. Although advantages on the attitude scale are in evidence for behaviour modification at all the follow-ups, these are on different dimensions within the scale.

A further issue that arises from considering the performance of these two regimes is that of making a choice between them, assuming of course that a choice is deemed necessary. If it is granted that they are roughly equally effective within the time span of the present follow-ups, albeit with slightly different 'trajectories' in outcome, a number of bases for making a choice can be suggested. First, these two regimes have differing philosophical bases, each with its respective appeal. Second, one is teacher-based, the other is not. This may be an issue of some importance given the possible conflict of opinion over teachers' assumption of responsibility for dealing with difficult and disruptive pupils. It should be noted that although in this study group counselling was conducted by social workers, it would not be unrealistic to expect that teachers or school counsellors could undertake similar kinds of involvement after appropriate training. A third question which gives an additional basis of choice is that of cost-effectiveness. It is not easy to make a decision on this basis by examining design and implementation in the present study since it is not clear, for example, how much or which parts of the social workers' training are necessary for the outcomes observed. Other important features are the numbers of field staff involved, with six group therapists in the present instance being compared with one behavioural consultant; the duration of treatment, with behaviour modification being twice the length of the group counselling programme; the possible <u>preventive</u> role of the treatments with future cost - reducing implications - the involvement of teachers in behaviour modification, for example, enables them to acquire skills which they could potentially apply with successive classes, or with emergent problems.

In selecting an appropriate method of intervention in the classroom, the choice is, of course, far wider than between the three models considered here. Given the range and complexity of tasks confronting anyone who grapples with classroom management issues, any one approach, such as behaviour modification, is likely to be limited in its contribution. As Brophy & Putnam (1979) point

out,.../

out, no single approach is adequate, and their suggestions as to what a comprehensive treatment of classroom management must attend to, are instructive, if rather daunting. It must attend to:

> 'relevant student characteristics and individual differences; preparation of an effective learning environment; organisation of instruction and support activities to maximise student engagement in productive tasks; development of group management during active instruction; techniques of motivating and shaping behaviour; techniques of resolving conflict and dealing with students' personal adjustment problems; and orchestration of all these elements into an internally consistent and effective system.' (p 215)

Other models can be seen as dealing with particular dimensions within such an array, and as adopting approaches somewhat different from those considered in the present study. A few of these will be briefly considered here. Glasser (1977), for example, advocates the use of classroom meetings between teachers and pupils in order to establish, monitor and change classroom rules and to deal with problems which arise. Such an approach clearly stresses co-operative strategies in dealing with management issues and also has implications for the development of pupil self-government. As with Glasser's model, Gordon's (1974) Teacher Effectiveness Training stresses the abandonment of power and authority on the part of teachers and favours the negotiation of problems through attempts at mutual meetings of needs. Redl's (1959) concept of life-space interviewing is in a similar mould, with its recommendations that teachers work together with pupils until each understands a problem incident and its meaning to the pupil, and until ways to avoid repetition of the problem are identified.

On a dimension dealing with the relative contribution of teachers and children to the resolution of problems, therefore, the above models differ considerably from the behaviour modification model examined in the present study. Somewhat closer to the behavioural position in its emphasis on the role of the teacher in classroom management is the work of Kounin (1970). He identified a number of features which were associated with success in general class management. The principal ones included the teacher's awareness of what is going on in all parts of the classroom ('withitness' in Kounin's terms) and ability to communicate this awareness to pupils; effective 'overlapping'..../

lapping', or the ability to do more than one thing at a time, which may involve handling individual requests or contacts without interrupting lesson flow; the ability to facilitate smooth transitions from one activity to another; and the capacity to sustain the attention of the whole class by a variety of 'group alerting' procedures.

These are but a few of the alternative approaches to the task of classroom intervention that one might adopt. Each of them involves a different set of assumptions concerning the nature of the problems to be dealt with and the manner in which they should be approached. The choice of an intervention model may be guided by a number of criteria, for example, the 'fit' between a model's underlying assumptions and one's own beliefs; the relative emphasis on teacher and pupil involvement; the available research evidence, and so on. More detailed accounts of models of classroom intervention are available in Dunkin & Biddle (1974) and Brophy & Putnam (1979).

(e) Treatment integrity

One of the issues which requires to be addressed in any treatment study is whether the intervention was actually implemented as planned. This is the notion of treatment integrity (Yeaton & Sechrest, 1981). If implementation is inadequate then the test of effectiveness is conducted on a weak or diluted form of treatment which may bear little comparison to the authentic model. In the present study, unfortunately, no clear answer is available to this question. First, feedback from the observer indicated wide variability amongst teachers in the extent to which teachers adhered to treatment recommendations. This tended to mirror the writer's own impressions of the teachers' response to the behavioural approach, ranging from obvious lipservice to great enthusiasm. These impressions, however, were not formalised in a manner which could afford a reliable estimate of their correlation with outcome. It was felt, along with the observer, that some 6 to 8 teachers (15-20%) failed to significantly alter their response. Second, observational data on teacher behaviour did not demonstrate changes in the predicted direction which would have been consistent with appropriate treatment implementation. However, it has been suggested.../

suggested elsewhere that this was an insensitive measure and may have consistently failed to register real changes. The significant differences between the two regimes with the consultation component, in favour of behaviour modification, would tend to suggest that the treatment was implemented successfully. Whether it was universally and consistently applied, on the other hand, remains unclear. There are a number of factors which could affect the quality of treatment implementation in a study such as the present. Some of these derive from the attempt to apply behavioural procedures within the context of an ordinary school, and these will be examined in the next section. One important source of potential difficulty that requires consideration here is the training consultation format adopted.

There are several features of this which may have served to weaken the intervention. First, the training itself may have been ineffective, at least for some teachers, so that habitual patterns of responding to children's behaviour were not altered in the desired direction. It was evident that a number of teachers, although apparently well motivated to change, experienced great difficulty in doing so. A frequent difficulty appeared to be the tendency to almost automatically make disapproving comments for minor misbehaviour, despite conscious attempts to guard against this. Thomas & Adams (1971) reported similar difficulties. It is possible that a more action-oriented approach, with guided rehearsal of procedures, may have facilitated changes.

Whatever the success of the early training phase, it is important thereafter to maintain changed teacher behaviour and help sustain their efforts. Although the model was explicitly set up as a time-limited one in terms of consultation, there was a continuing feeling on the writer's part that the time available was not sufficient for the discussion, planning and general support necessary with the teachers. With teacher availability limited to break, lunch or non-teaching time, there was pressure to make maximum use of the time given which did not always make for the best conditions for dealing with sensitive issues. It often felt as if some issues had been left half dealt with. It was also difficult at times to avoid a sense of encroaching on teachers' free.../

free time - and with some teachers, not surprisingly, the time seemed reluctantly given. At times there was considerable strain for the writer in maintaining the level of contact required, especially when the implementation phase overlapped across several schools. Much travelling was involved both in conducting consultations and assessments, and with time having to be seized whenever it was available in teachers' timetables, a great deal of shuttling backwards and forwards was necessary between schools which might be up to ten miles apart. (The 80 hours spent on seminars and consultation did not include this travelling time.)

There were several reasons, therefore, why the consultation arrangements were less than ideal, and perhaps the large numbers of children and teachers involved, and their distribution prevented a fair trial of even a time-limited model. During the process of training and consultation, many issues arose in connection with teachers' response which seemed to have important implications for treatment implemention. With a large group of teachers who have not been hand-picked for the task, it is inevitable that a number will be poorly motivated. The sources of poor motivation in this context are many. The next section will address some of these problems along with others which may accompany interventions in regular school settings. In highlighting these problems, it is implicitly recognised that they would require close attention in a future intervention along the lines of the present one.

(ii) <u>Problems in Implementing Behavioural Procedures in Regular School Settings</u>

The preceding section has pointed to some ways in which training of teachers and ongoing consultation proved difficult to arrange within and around working practices in regular school settings. Some of these difficulties may have affected the quality of treatment implementation, and they have to be seen alongside a number of other factors which may influence treatment delivery in such settings. These include the various demands and influences which operate in the classroom, teachers' own working philosophies, and pressures which may arise within the school as a social setting when a new intervention model is introduced.

Many of the issues to be discussed here came to light in the course of the project and through experience of working with the teachers in the schools. In retrospect, if programmes such as the present are to be applied successfully in regular settings, issues such as these may require almost as much attention as the intricacies of the treatment application itself.

(a) <u>The demands of the classroom</u>

The first source of difficulties to be examined requires consideration of the demands made on a teacher through the introduction of a new treatment model, and howthese demands exist alongside many other influences and pressures in the classroom. In the present study, where intervention or 'treatment' were very much secondary concerns, with teaching the foremost priority, it is important to consider the numerous influences on teacher behaviour and how these might have affected treatment implementation. The classroom is undoubtedly a complex social situation, with many demands being made routinely on the teacher. It should not be assumed that the set of demands associated with treatment procedures somehow neutralises these other influences by becoming paramount in the teacher's awareness and dominating all else.

What are some of the features of classroom life and teacher behaviour which may affect treatment delivery? The following is a selective outline. First, the classroom is a busy place. In a single day, a teacher may engage in more than one thousand interpersonal exchanges with pupils (Jackson, 1968). A study of four teachers of 11-12 year-olds found that they averaged 80 initiations with individual pupils each hour (Jackson & Lahaderne, 1967). In addition to initiating exchanges, teachers will move about, explaining here, responding to a question there, monitoring and managing a variety of events occurring simultaneously, anticipating others. In the midst of such constant and varied pressures, and the need to respond to present and immediate demands, it is easy to see how teachers may experience more than a little difficulty in planfully and systematically applying behaviour modification procedures. Good & Brophy (1978) argue indeed that teachers' absorption in the demands of .../

of the classroom militates against the development of a clear awareness of their own behaviour. Such awareness may well be a prerequisite for change in teaching practices.

Within this prevailing busyness, it appears that teachers' attention may be attracted and drawn in certain ways which could well compete or conflict with the requirements for teacher attention specified within a behavioural programme. Thus, boys seem to be much more salient to the teacher and receive much more criticism than females in both primary and secondary classrooms (Lippitt & Gold, 1959; Jackson & Lahaderne, 1967), and low achievers receive less response opportunity (Mendoza et al., 1972) and less frequent and favourable contact (Good et al., 1972) than their high-achieving counterparts. The quality of interaction with pupils may also vary with achievement level. Studies by Rowe (1969) and Brophy & Good (1970) suggested that teachers waited significantly longer for more able pupils than for less able ones before giving the answer or calling on another pupil. With abler pupils they might repeat the question, provide a clue or ask a new question, and expect performance, but with low achievers they were more likely to give up or accept minimal performance. An important determinant of interactions in this context, therefore, is the quality of feedback the teacher receives from the child in the exchange, or put another way, the extent to which the teacher experiences the exchange as rewarding or frustrating.

A further influence on interaction and communication appears to be the manner in which pupils are grouped. Certain seating arrangements, for example, may channel the teacher's attention in particular ways. Adams & Biddle (1970) described an 'action zone' in the classroom, in which children seated in the middle rows and towards the front tend to receive more teacher attention and also receive more opportunity to interact with the teacher. A passive, low-achieving child seated outside of this zone may find that little attention comes his way. In order to systematically reinforce such a child for desired responses, therefore, a teacher may have to consciously and repeatedly break his own customary patterns of behaving. Such 'action zones' presumably enable the.../

the teacher to reduce the demands of attending to and dealing with the class group, so that interactions are most frequent with those children within easiest reach or who are the most accessible.

These, then, are a number of influences which have been identified as operating on teacher behaviour. There are probably many others, including countless idiosyncratic ones for individual teachers. What is being suggested here is that influences such as these will have co-existed with the more identifiable 'influence' of intervention strategies. Whether they hindered or potentiated these strategies is not known since their role was not specifically addressed within the experimental design. For some teachers, it is quite possible that these factors contributed to what may simplistically be called 'poor teacher motivation' for change. It must also be acknowledged that the writer had little awareness of many of these processes until the project was well under way and some of the complexity of classroom life was becoming apparent. It is also to be noted that most of the above findings emerged from work carried out not within the context of behaviour modification but within the body of eduational research concerned with classroom observation. There has been little interpenetration of knowledge between these two fields, despite the numerous ways in which one could enrich the other. Certainly from the perspective of someone introducing a treatment or management model into the classroom, knowledge of its hidden complexities and undercurrents would undoubtedly be of assistance in clarifying the network of influences within which helping efforts are to be embedded. The impact of behavioural procedures may well be improved by drawing on the findings of other models of educational enquiry.

(b) <u>Teachers' acceptance of behavioural approaches</u>

A second consideration in the introduction of behavioural programmes into regular settings is the degree of acceptance of the philosophy of behaviour modification amongst the body of teachers.

It was evident with the present group that a number of teachers had objections to the approach although this did not always mean that they were unwilling to experiment with the procedures. Teachers, not surprisingly,

believe.../

believe in particular sets of educational theories and hold certain preconceptions in the same way that mental health workers do. It is therefore understandable that teachers' own personal and professional investments could be at variance with the framework of assumptions offered by behaviour modification. Although within the research context there was pressure to pursue the application of procedures, it seems important to recognise that the behavioural approach is but one avenue to seeking change and it is essential to respect someone else's guiding philosophy, even if contrary to one's own. This was an important consideration, not just at the outset but throughout discussions with the teachers.

The criticisms that teachers themselves offer of behaviour modification often serve to illustrate the gap between their own beliefs and behavioural principles. Perceptions of reinforcement as bribery are quite common. This issue has been examined in an earlier chapter. An associated concern is the emphasis on positive reinforcement. It is quite clear that, for some teachers, the underlying issue here is a fear that children will view them as 'going soft' and losing their 'firmness'. For others, their objections may reflect a recognition of the considerable changes required in their own behaviour if the desired emphasis on positive reinforcement is to be achieved.

A criticism that has important implications for the viability of behaviour modification is the one where responsibility for problem behaviour and its management is disowned by the teacher. This criticism is often voiced not as a specific criticism of behaviour modification per se, but as a more general statement. Thus, some teachers regard a special class within the school, or a special setting outside it as the proper province for treatment, rather than their own classroom. Other teachers view children's problems as centred within the child, deeply and permanently ingrained, or irreversibly tied to adverse family or social factors. Manipulation of environmental influences within the school setting, in line with behavioural principles, is therefore irrelevant and inconsequential. In one study in which teachers were asked to explain why their pupils presented problems, only four of 74 teachers perceived themselves as contributing to the problem (Good et al., 1969).

Nash ..../

Nash (1976) suggests somewhat cynically that teachers have a vested interest in maintaining the assumption that the important determinants of educational and behavioural functioning are to be found outside the classroom. Thus, of a child who does well, it can be said: 'We taught him successfully', and of a child who fails, it can be said: 'Terrible background, can't expect anything else.' Nash does not appear to concede that taking credit for success but not the blame for failure is very much a widespread human failing, not one specific to teachers.

Criticisms such as these do need serious consideration and discussion as they represent threats to the adequate implementation of a behavioural model. The recent policy shifts towards retaining problem children within the mainstream of ordinary schooling provides an impetus for a re-analysis of issues of responsibility, so that behaviour modification can be advanced as one helpful strand in the coping responses that some teachers are necessarily going to have to develop. With those teachers who are concerned that their position as effective controllers of discipline might be undermined by what they see as an excessively positive approach, it is important to explore issues such as the ways in which positive approaches are not incompatible with 'firmness', the possibly undesirable side effects of punishment, and the need for a balance between positive and negative forms of management.

(c) <u>The social context of intervention</u>

A third set of factors with implications for quality of teacher response and acceptance of new methods requires consideration of the teacher's place in the social grouping of teachers within a particular school. With the experience gained in the programme, it seems important to recognise that the response which is usually referred to as 'resistance' - to change, innovation, different practices - is not necessarily just an <u>individual</u> response to feeling personally at odds with the changes required, feeling imposed on, not being consulted properly and so on. There is an important sense in which resistance reflects the individual's place within the group of teachers or within the organisational system of which he is part. Any change undertaken by an individual has.../

has implications for others within the group or system. He may be perceived differently or feel that he is perceived differently, role or status relationships may be affected, and so on. In three schools in the study a number of teachers mentioned their fear of openly voicing support for the behavioural project in the presence of senior staff, reflecting the fear of being in some way cut off from the group. In such circumstances, reluctance to pursue change can be seen as very much a self-protective response. A similar element of 'face preservation' can be seen in the reluctance of some teachers in verbalising their classroom control difficulties in group/seminar settings, although being more than willing to do so in individual consultations. One teacher described a sense of hostility towards her and her involvement when a colleague asked her rather disdainfully: 'Are you going to be one of the carrot-danglers, then?'

The pressures on those who attempt to initiate change within an organisational group is illustrated by Moss & Childs (1981). Writing in a British context, they report how the attempts of a small number of teachers in a secondary school to dispense with a lunchtime detention system, as part of a token economy scheme, were eventually undermined by another influential group of staff who wished to retain it. They comment that 'here was a sub-system trying to influence the larger system but within which were locked other conflicting sub-systems that operated powerfully to maintain the "status quo" rather than face the threat of change.' (p 97)

Recognition of these difficulties might suggest that it would be helpful to organise support systems of staff who are favourably identified with the project, which would reduce the sense of isolation. There seems to be much truth in Sarason's (1971) comment that 'teaching is a lonely profession'. It was frequently remarked by teachers in consultation sessions that the behavioural programme constituted the first constructive offer of help in dealing with classroom problems. Little or none seemed to be forthcoming from colleagues. Hargreaves (1978) has commented on the stress suffered by new teachers and how this is heightened by professional isolation. This arises.../

arises not because teachers are unfriendly but because each leaves the other to solve his own problems:

'The informal rule among teachers is that every teacher is king in his own classroom. Any interference which threatens that autonomy is to imply one's own incompetence and to volunteer help to another is to impute incompetence. The result is that teachers bear their stress in painful isolation. It attacks the heart of the teacher, both physically and metaphorically.' (p.541)

The lack of recognition given to difficulties such as these would suggest that the creation of special staff support groups to deal with such issues, while potentially helpful, would be no easy task. With current trends in educational policy towards the retention of 'difficult' and maladjusted children within ordinary schools rather than placing them in alternative 'special' settings, it is likely that increased attention will be given to ways in which staff may be supported in meeting additional demands on their coping capacities. Hanko (1981) has offered a useful account of how group consultancy support can play a valuable role in ordinary schools. Of particular relevance to the present discussion is her conclusion that:

'It (is) possible to develop in consultancy support groups, a professional commitment and maturity within which unpreventable occasional failure could be accepted without excessive frustration, guilt feelings, or defensiveness, so that colleaguial support could be asked for, or offered, as part of one's professionalism. This is in contrast to the still far too frequent isolation of teachers with problems within schools, or the defensive in-group rallying of staffs maintaining that "there are no problems in their schools", with only "fire-brigade" or "launderette" forms of recourse to outside experts, and abdication to them, much regretted by the experts themselves.' (p.27)

Within an intervention model such as the present one, discussions with pairs of teachers or with small groups would be one way of developing supports within the colleague group. Although the need for organising more substantial support for teachers was recognised during the running of the programme, limited time and the difficulty of arranging group meetings within the constraints of the timetable effectively militated against this. In a general way, more attention could also have been given to discussion of the project with uninvolved staff <u>during</u> its implementation, so that the 'public relations' and smoothing-the-way exercises were not confined to the very introductory stage. The involvement of senior staff is obviously important here, in that they hold authority and influence, but significant power and status relationships can also exist <u>within</u> more junior groups of staff.

Interventions which do not give due regard to these facets of the system within which programmes are located are open to the criticism of operating under the belief that organisational change can be achieved by changing individuals. For Katz & Kahn (1966) such an assumption 'is an oversimplification which neglects the interrelationships of people in an organisational structure...' Georgiades & Phillimore (1975) view this kind of approach as:

'....the myth of the hero-innovator: the idea that you can produce, by training, a knight in shining armour who, loins girded with new technology and beliefs, will assault his organisational fortress and institute changes both in himself and others at a stroke. Such a view is ingenuous. The fact of the matter is that organisations such as schools and hospitals will, like dragons, eat hero-innovators for breakfast.' (p. 315)

A number of writers (e.g. Repucci & Saunders, 1974; Colman, 1975) have pointed out that workers in the behaviour modification tradition have ignored or not dealt with the complexities of such social settings. This occurs partly because the problems do not relate directly to theoretical issues within their own field, and partly because in highly controlled laboratory or special research situations, such difficulties are inconspicuous or absent. Repucci (1977) suggests that failure to appreciate the importance of such issues and the more general relationships between community psychology and behaviour modification may be a primary reason why demonstration projects so often fail when efforts are made to transfer them beyond the research clinic-laboratory or the isolated demonstration classroom.

Knowledge acquired within the operant framework may well not be adequate to the required analysis, in that, as Colman (1975) puts it:

'Operant techniques have always presupposed the ability to gain control of the critical reinforcing contingencies influencing the behaviour to be modified.' He goes on: 'Token programmes.....rely on the ability to manipulate the rewards in the system including the social rewards such as attention and praise, which are part of the behaviour of personnel. Organisational design, however, relates as much or more to staff and administrative groups as it does to patients.' (p. 336)

If the theory of change is therefore to be appropriate to and mirror the complexity of social settings,then knowledge in other areas must be drawn on. Of obvious relevance is work within the areas of attitude change and organisational psychology. It is also encouraging that analyses of school-based.../

based problems are increasingly drawing on systems theory (e.g. Burden, 1981). Consideration of such aspects of the social context suggest that in regular school settings, where interventions are conducted with available rather than highly selected teachers, such difficulties may be the norm. Stringent implementation of treatment therefore may not be as readily guaranteed as it can be in the more rarified atmosphere of laboratory or specially established experimental classrooms. Thorough-going implementation of behavioural procedures has, of course, been achieved in regular school settings, but this has commonly involved small numbers of teachers or specially selected ones. The relative paucity of research interventions in regular settings suggests that much work remains to be done in identifying the reach and limits of behavioural procedures - with which children, in which settings, with which teachers - before the problems of large-scale field applications are resolved.

(iii) Critical Views of Behaviour Modification

The foregoing sections have considered some of the problems which may require to be dealt with in introducing behaviour modification into schools. It is also necessary to go beyond this essentially narrow context of concern to consider some of the difficulties facing behaviour modification more widely. This requires consideration both of some of the criticisms levelled against the approach and of the model of inquiry within which it is embedded. This will hopefully provide a wider perspective from which to view the present investigation, as well as a reminder that behaviour modification is but one route to intervention and inquiry.

(a) The focus on behaviour

A frequent criticism of behaviour modification revolves around its focus on behaviour and the way behaviour is conceived in topographical terms (e.g. Pring, 1981). What is objected to here is the limiting and restricting separation of physical behaviour from that which ultimately gives it its meaning - the mental world of beliefs, desires, intentions, feelings and so on. Implied in this separation by behaviourists is the notion that such unobservable concomitants of behaviour are not necessary in the explanation of behaviour. This.../

This is the basis of the radical behaviourist position which seeks an understanding of behaviour through the functional relationships established between behaviour and observable environmental events. This stance has been hotly disputed, on a number of counts. Thus, the importance of internal events has been emphasised by investigators who stress the contributory role of awareness of experimental contingencies in human learning. In the well known demonstrations of verbal conditioning (Greenspoon, 1962) it was claimed that the subjects were unaware that the experimenter was providing reinforcement, and that learning effects were automatic and unconscious. Subsequent investigations demonstrated that only those subjects who became aware of the response-reinforcement contingency actually showed learning effects (De Nike, 1964; Spielberger & De Nike, 1966). Similar findings have been reported for other areas of learning (Murray & Jacobson, 1978). The contribution of awareness is still disputed (e.g. Brewer, 1974; Dulany, 1974), although this is belied by the provocative stance taken by Bandura (1974) who declared that, 'contrary to popular belief, the fabled reflexive condition in humans is largely a myth'. At the very least one can justifiably conclude that awareness facilitates learning.

In situations where reinforcement is being made available it has been argued that subjects will try to provide interpretations or construct hypotheses as to what reinforcement means. For example, it has been shown that if a subject believes he is being reinforced on one particular schedule, but is actually being reinforced on another, he may behave in accordance with his beliefs instead of in accordance with the real contingencies (Baron et al., 1969). Similarly, the effects of conditioning procedures may depend on whether the subject attributes the occurrence of reinforcement to his own skill and efforts or to some outside agency or mere chance (Rutter et al., 1961). For Bandura (1977), response consequences impart information and serve as motivators through their incentive value. As he argues, people not only perform responses but also notice the effects they produce in the course of learning. Reinforcing consequences inform them what they must do to gain beneficial outcomes and to avoid punishing ones. Anticipatory capacities.../

capacities enable humans to be motivated by prospective consequences. Thus, the sense that is made of past experiences creates expectations that certain actions will bring benefits, that others will have no appreciable effects and that still others may have negative outcomes. By representing future outcomes symbolically, people can convert future consequences into current motivators of behaviour. So in Bandura's terms, most actions are largely under anticipatory control. For Franks & Wilson (1979) there is no question that Bandura's 'additions' to basic stimulus-response conditioning accounts have contributed to improved understanding and outcome.

Such a perspective on motivation is a long way removed from the radical behaviourist stance which would view the issue in terms of questions to do with the contingencies and relationships associated with the prevailing schedules of reinforcement, and the probabilities of responding associated with a given set of discriminative stimuli. The recent growth and current popularity of 'cognitive-behavioural' interventions (e.g. Kendall & Hollon, 1979) suggest that the prevailing Zeitgeist is now identified with a concern for incorporating mediational processes, but within a context of concern for methodological rigour. These developments are consistent with a view of the scientific enterprise which regards unobservable events as legitimate objects of concern. This, as we have seen, is an unacceptable position for radical behaviourists, yet it is apparent that even in the physical sciences adopted as a model by behaviourism, unobservable phenomena <u>are</u> dealt with and there is an abundance of inference (Erwin, 1978).

(b) <u>The wider implications of explanatory models</u>

For a critic such as Pring (1981), the adoption of explanatory approaches which exclude any reference to the understanding and intentions of the agent, which normally enter into an account or description of what is happening, will have far-reaching implications. Not only will explanations of behaviour be depleted by failing to consider how certain motives or plans may have coloured the choices and decisions made, but the terms of reference of the explanatory system will become narrowed and constricted. Pring argues that:

'such..../

'such a restriction of meaning will affect the predictions of what will happen because what is to be "observed" (if the predictions are correct) will be described differently. One generally perceives only what is anticipated in the concepts one has - restrict those concepts (as programmes of behaviour modification would have us do) and one restricts these perceptions.' (p.66)

The 'restrictiveness' or self-limiting nature of behaviour modification in this sense undoubtedly grew out of its identification with a model of science which assumed a detached and objective stance towards its areas of concern. Rothkopf (1968) sees the problem as the basic one of the transfer from the laboratory to the real world. In the laboratory, precision, objectivity and control are possible, but when the moment of transfer to the outside world arrives, does one alter the conceptual system so that it fits the existing realities outside, or does one try to remake the outside world like the laboratory? For Rothkopf, the behaviourists opted for the latter, and the reductionists' 'simplicity filter'. In referring to laboratory experiments during one of the training seminars, one of the teachers commented that: 'All this is really Utopian - it's irrelevant to real classroom situations'.

For many critics, one of the most objectionable features of behaviour modification is not its construction of the outside world in the manner of a laboratory, but rather its derivation from experimentation in the animal laboratory. Such roots may represent another influence towards the over-simplification of explanations of human behaviour. In addressing the shift from basic laboratory settings, Davey (1981) outlines an idealized sequence of steps for the application of conditioning principles for therapeutic use. His account indicates how, even if one accepts animal experimentation as a valid starting point, the transposition to humans may not be well-ordered and controlled. First, the principles are established in animal studies. Second, laboratory studies should ensure that these principles apply to humans. Third, the principles are adapted into a technique to cope with the practicalities of the therapy situation. Davey argues that the second step is frequently bypassed and that the conditioning procedure carried out in the third step is often mutated to a point where it is no longer an accurate representation of the original principle. Others have pointed to how far.../

far removed procedures applied with humans may be from those established in the animal laboratory. Thus, Sajwaj & Dillon (1977) point to the confusing complexity of an adult's differential attention - a procedure that is often viewed as a 'simple' reinforcement technique. Similarly, Kazdin (1977b) has discussed the many social factors which complicate interpretation of the administration of tokens within a token economy, so that it cannot be viewed solely as a straight-forward application of a unitary reinforcement. In the earlier review of operant conditioning principles, the need was noted for precise temporal relationships between reinforcement and the behaviour to be strengthened. Davey (1981) indicates that within the context of self-control studies, much human learning takes place without observing this contingency. He additionally reviews evidence which suggests that, in human subjects, non-reinforcement does not always produce extinction, and that different schedules of reinforcement do not always produce predictable rates of responding.

This divergence between animals' and human behaviour in their respective experimental settings suggests that, even though basic conditioning principles may be applied to some degree in explaining human behaviour, supplementary principles may be needed to explain what we observe. Greater attention may need to be paid to distinctively human capacities in these situations, such as awareness of reinforcement contingencies and constructive attempts to form an understanding of the processes involved.

Criticisms such as the ones considered above are not arguments against the use of behaviour modification programmes, but pointers to the potentially restricted nature of their application. It is perfectly legitimate to explore the contribution of such a 'restricted' model to the solution of children's problems in school, as the present study has done, and it is a matter for further investigation whether other 'less restricted' or alternative models have more to offer in this particular area of concern. Relatedly, it may be accepted that behavioural procedures 'work' in applied settings, but it is justifiable to ask <u>why</u> they are effective. In a therapeutic context, it is important to ask whether a radical behavioural account of.../

of effectiveness is adequate or whether greater treatment benefits would flow from incorporating explanatory accounts which lie outside its framework.

There is implicit in the adoption of any explanatory account a model of man, and the need to maintain a perspective on the behavioural model - and indeed any other - as potentially limiting, cannot be overstated. Models have the capacity of functioning as self-fulfilling prophecies. As Quicke (1975) points out, in a critique of behaviour modification in education, the notion that man can be represented as a machine, with behaviour composed of learned and associated units, has important implications for our conception of human beings:

'Think of people as machines for long enough and they become machines. Emphasise that the content of psychology is that which is observable and measurable and eventually the trivial which is the only aspect of human functioning amenable to direct observation and measurement, the hard data, becomes the most important data, then the only valid data, and man himself is trivialised.' (p.10)

Although what is regarded as 'trivial' and 'valid' is open to debate, and requires to be considered in the light of what behaviour modification has achieved, Quicke's essential point remains an important one. Models of understanding have the knack of creating and producing their own reality. This same point is also relevant to discussion dealing with criticisms that behavioural procedures serve as instruments of manipulative control, often serving the needs of those in authority and helping to sustain the existing order. (It is to be noted here that even non-coercive and psychologically appealing procedures, such as social reinforcement, still constitute imposed control. Docking, 1980) The well publicised review by Winett & Winkler (1972) is very much in this vein. These authors sharply criticised behavioural classroom programmes whose sole goal was to teach children to 'be still, be quiet, be docile!' They concluded that the pupil produced by these programmes was one 'who stays glued to his seat and desk all day, continually looks at his teacher or his text/workbook, does not talk to or in fact look at other children, does not talk unless asked to by the teacher, hopefully does not laugh or sing (at the wrong time), and assuredly passes silently in..../

in the halls.' (p. 501) These authors argue that deviance may be beneficial to society, in that it indicates ways in which the social system is itself inappropriate and in need of modification.

A typical rejoinder to criticisms such as those of Winett & Winkler, which imply control or questionable objectives is that the techniques themselves are ethically neutral and do not convey which ends should be sought, only the means of achieving certain ends. Skinner (1972) himself, for example, stated that 'no theory changes what it is a theory about'.(p.215) However, it is here that the earlier argument, that models of understanding create their own reality, needs to be restated. As Dearden (1981) points out:

'People are importantly changed by the pictures which they form of themselves and which others form of them. The pictures embodied in educational practices, in different religions and in political and economic theories are at least partially explanatory of why people regard themselves in the ways that they do, for example as autonomous, as immortal souls seeking salvation, as part of a socialist brotherhood working for the common good, as engaged in individualistic economic competition and so on....with people, belief is apt to beget reality.' (p. 117)

However, as Dearden himself concedes, there is still some justification for focussing criticism on certain behaviour modifiers for the slant which their own personal values or attitudes give to these techniques. In the classroom setting, it probably is not surprising that where there is no consistent agreement as to which classes of behaviour are important or essential for children to have, teachers often select as targets those behaviours that would be most reinforcing to them, by helping them to maintain control over their classrooms (Stoltz, 1978). It is interesting in this respect that few teachers questioned the aims of the procedures in the present study.

It would seem necessary therefore, that much consideration needs to be given to the array of objectives that theoretically could be pursued within the classroom. Is an adequate array of objectives being considered? Whose interests are being served? Even if objectives are agreed upon by all parties concerned, however, it may still not be appropriate to pursue them via behavioural procedures, since the ends may not justify the means. Extrinsic reinforcement, for example, using powerful concrete reinforcers, may.../

may be the most 'effective' way of accelerating academic output, but many would object to such an arrangement where motivation is separated off from the curriculum itself (Docking, 1980).

(c) A limited model of science?

What is implicit in some of the foregoing criticisms is that the positivist model of science out of which behaviour modification has grown is quite inadequate to deal with human behaviour and action. This model, which has adopted the detached stance of the natural sciences and its methods of controlled observation and experimentation, involves the assumption that there is a similarity between human, social life, and the natural world, which makes a transposition of methods valid. Research within this model is characterised by Shipman (1981) as 'top-down', imposed by the researcher on to the respondents, and typically the researcher controls the conditions under which responses are observed or measured. The model involves imposing the stimuli which produce these responses, and the terms in which responses are to be interpreted. The interest in the respondents is in their behaviour and its fit into established or anticipated models. The respondent contributes only reactions to pre-determined stimuli and there is no accommodation for spontaneous behaviour. Yet, as Shipman goes on to argue, 'humans are unsuitable for study by methods designed for studying the unthinking'. (p.133) To impose preconstructed categories on social life in order to study it may merely ensure its distortion. Given that human beings are capable of making sense of, and interpreting their world, these attributes must be given a central role if the model of understanding is to be adequate. As Shipman puts it: 'The natural scientist can construct, control and define the phenomena he intends to study, because the subject matter is not providing any alternative set of constructions. But the social scientist faces humans who have already interpreted their world, will continue to give it meaning as they are investigated, will also interpret the actions of the researchers and are capable of coming to their own conclusions about the meaning of events'. (p.133)

For Shipman, an alternative, 'bottom-up' model of research, involves a radical revision:

'The focus should be on the actor and the way he perceives, interprets and assesses the social world around him. The conventional researcher's role of detached observer is useless. The appropriate position for research is from the actor's viewpoint. The researcher has to take on the role of the actor. To remain detached, to attempt to preserve objectivity, means imposing on the actor's own perceptions and consequently ensuring their distortion'. (p.136)

An example of this alternative framework in which humans are given full scope as planning, reasoning persons who are capable of giving valid accounts of their interactions is the study by Rosser & Harre (1976), which points to the disparity between the motives for misbehaviour that teachers ascribe to pupils and the motives given by the pupils themselves. Some responses which teachers regard as insulting may, for the children concerned, involve the protection of self-esteem. Interviews with pupils suggested that pupils who withdraw by going silent and unresponsive do not always see themselves as returning contempt for contempt but rather as using a policy of 'equilibration' to restore themselves as human beings. This shows the contrast between teachers' and pupils' accounts or explanations of the same episode. Other examples are the work of Hargreaves et al. (1975) on the attribution of deviance to pupils, and Woods' (1976) analysis of the functions of 'having a laugh' in the classroom. Woods (1980) has recently edited two books 'Pupil Strategies' and 'Teacher Strategies' devoted to a collection of work from the 'bottom up' standpoint, which serves to illustrate the expansion of work in this area. Some of the wider implications of these recent developments are considered by Reason & Rowan (1981) under the heading of 'new paradigm' research.

There is little question that these new approaches can enrich explanations and understandings of behaviour in the classroom. They are not, however, without their own difficulties. If the intention is to convey faithfully the perspectives of the participants in an interaction - say a teacher and a pupil - how may a researcher's own concepts be excluded, or at least be identified as such? This is especially difficult where a summary of participants' accounts is to be made, or where some interpretation is offered. How may differing accounts be weighted? It would seem essential, given the theoretical standpoint.../

standpoint adopted, that accounts be gathered from principal participants and cross-checked against one another. This would bear on questions of the reliability of the information. Yet, as pointed out by Shipman (1981), a study such as Marsh, Rosser & Harré's (1978) 'Rules of Disorder' omitted any account from the teacher in an examination of classroom relations. Other questions arise in connection with the role of the researcher. For example, how is the problem of his own influence on responses dealt with or controlled? How are non-verbal as well as verbal cues from the researcher handled? There are therefore many problems of credibility to be solved within this new framework, perhaps, as the cynics might suggest, as many as in the traditional framework that it is attempting to replace.

(d) Taking account of new perspectives on research inquiry

It is of interest at this point to consider ways in which the present study could have moved closer to this more co-operative model of inquiry, in which the client's contribution to planning and decision-making is taken more fully into account. It is necessary, however, to distinguish between the different interpretations of 'client'. In the discussion so far, the reference has been to the children whose behaviour it was sought to change. There is also a sense, however, in which the teachers were 'clients' within the training and consultation exercises, since change in their behaviour was regarded as a prerequisite to changing children's behaviour. In addition, in the sense that the research project was brought to the school, and sought to change certain components within its system, the school organisation could be regarded as having some 'client' status. Beginning with the role of the child in the behavioural approach, and the implicit model of the child that it conveys, it is evident that this was quite limited and constrained. Thus, with the few exceptions with whom behavioural contracts were developed, the children were not consulted about the goals and objectives of the programme, these being established principallywith the teaching staff, and taking their views into account. A more co-operative approach may have engaged the.../

the children in discussion of issues such as classroom rules, how they evolve, how they differ across teachers, perceptions of different teachers, and so on.

It is possible, on the above perspective, to see the behavioural approach as viewing the child more as an object rather than as a subject, as someone to whom things are done, or as someone attended to in terms of his behaviours. In this respect, Quicke's (1975) suggestion that models of understanding generate their own realities seems especially relevant. It was also apparent that only very small snippets of the child's life - even within the school and excluding his life outside it - were taken into account. Thus, his time in one or two subjects, or with one or two teachers was examined. In a comprehensive school, however, a child may be exposed to around fifteen teachers in the course of a week. It is possible that our understanding of a child in only one lesson would be enriched by having a wider knowledge of what his life is like throughout the school day or week. Following children through their variety of experiences, across teachers and classrooms, and examining the sense they make of these changes, might be a useful alternative approach.

It is to be recognised, of course, that proceeding along such lines as suggested above would make for a markedly different kind of study from the present one, and, in particular, would make for a different set of introductory questions, goals and objectives in framing the inquiry. Turning to teachers and staff within the school system, it is likely that they could have been involved differently in the research process, and in ways which could possibly have increased the impact of treatment interventions. Unfortunately, the practical demands associated with treatment implementation had the effect of focussing attention too closely on the specifics of treatment, and a number of processes which had an important bearing on implementation were detected only well into the treatment phase, or with hindsight. It is perhaps inevitable that in this form of research, (as in many others), processes and influences outside the ambit of one's own theory are relatively

neglected.../

neglected, so that it is hard to imagine that one's next attempt, or a replication, would be conducted in the same fashion. Once such 'peripheral' processes are recognised, and their relevance acknowledged, they begin to influence aims, objectives and methodology.

There are a number of ways in which the approach of viewing the teachers or school system as 'clients' fell into a 'top-down' rather than a 'bottom-up' research model, in Shipman's (1981) terms. First, and most importantly, the school staff did not ask for the research project and its treatment methods to be brought into the schools. Although permission was granted for entry into the school, and staff offered their services in the various interventions, school staff did not invite the project nor where they involved in its detailed planning and organisation. Their absence from participation at this level, along with the fact that the research ideas were conceived not by educational but by 'mental health' personnel, may have had an important bearing on their investment in the programme. Where an intervention is <u>brought</u> to an organisation rather than <u>invited</u> into it, the best course may be to seek out people who can offer supports, rather than seeking change across the board. As Georgiades & Phillimore (1975) put it:

> 'The manager of the change effort should work with the forces within the organisation which are supportive of change and improvement rather than working against those which are defensive and resistant to change. It is far better to find someone who wants to help and wants you to work with her or him than it is to try and convince sceptics that they need your assistance.' (p.316)

The research design, however, did not allow for such selectivity of staff, calling, as it did, for randomization by class. To an extent, as well, it was necessary to test out the feasibility of conducting treatment interventions in this fashion since part of the intention was to experiment in natural or regular rather than highly controlled conditions. Georgiades & Phillimore also suggest that the path of least organisational resistance be followed, rather than confronting the resistance. This seems somewhat short-sighted advice, for if the intervention is to survive within the organisation, surely elements of resistance have to be dealt with? Not to do so is tantamount to denying or avoiding the reality of existing social forces and...../

and relationships. It was suggested earlier that 'resistance' may in some respects reflect the pressures that an individual may feel within an organisation - pressures, for example, to oppose certain innovations or to maintain the status quo. It would seem important for a researcher to be sensitive to such issues, and if possible, evolve ways of dealing with them.

A second way in which teachers could have been given a voice was in the selection of children for inclusion in treatment groups. Although hardly any children who were by common consent 'problem children' were missed by the screen, and although some teachers were involved in producing behaviour ratings, a number of teachers indicated that they would have preferred a freer nomination system. An associated difficulty was that from time to time teachers asked for advice or recommendations concerning children outside the treatment group. In the normal course of events, such discussion would be freely undertaken and would play an important part in continuing co-operation and an atmosphere of working together. Under the existing conditions, however, the limited time available for consultation or the need to avoid contamination of control cases meant that such discussion was minimal or not undertaken at all.

Under circumstances of staff involvement such as the above, the phenomenon usually referred to as 'resistance' on their part is not unlikely. It is in this light a natural reaction, all the more so when an uninvited research project is interpreted as in some way implying that they are 'not doing it right'. This implication is hard to avoid when teachers' customary ways of doing things are to be changed, however much stress is put on the 'purely experimental' or 'exploratory' nature of the exercise.

It was suggested earlier that staff support groups might make a useful contribution in opening out discussion of teachers' problems in the classroom. Such groups might also provide a forum for dealing with interpersonal issues which arise within the school's social system. After one rather emotionally charged seminar in which much friction was evident between sub-groups of staff, one of the teachers commented: "What we need isn't talks about behaviour modification..../

modification but sessions on interpersonal relations. Without that, it's as if you're wading through porridge."

(iv) Methodological Implications

This section will address a number of issues related to the design and methodology of the present study which require attention. Some of these are viewed as positive features, others make for problems in interpretation and analysis of findings: they all have implications for future research.

(a) Assessment procedures in evaluating outcome

Mention has already been made of the merits of including multiple measures in initial assessment and in monitoring of change, and conducting assessment over an extended follow-up period. The present study has been rather exceptional, in these terms, amongst behavioural studies in educational settings. Its findings would tend to suggest that the use of multiple measures and extended follow-up is well justified. Much important information concerning the extent and duration of changes may have been lost in previous studies simply because theoretical preconceptions suggested that anything other than limited - usually observational - measures conducted within a short time span was likely to be unproductive. (Theoretical preconceptions in any field of enquiry will always, of course, result in a narrowing of the field of concern, through the particular focus that is adopted.) While it is recommended, however, that future studies should pay heed to these requirements, it is to be noted that considerable work remains to be done firstly in the explanation of correlated changes and, secondly, in relating changes in the long term to intervention procedures that are remote in time.

(b) Comparative focus

Another apparently well justified feature of the study design is its attempt to compare several treatments. Such a design is a considerable advance over treatment versus non-treatment comparisons in terms of its ability to extend explanations of change beyond simple placebo or attentional factors. An intriguing feature of the present comparison, however, is that far.../

far from highlighting the divergent contributions of different treatments, it has pointed to the similarity in impact of two theoretically divergent models. As already noted, the inclusion of measures in the assessment battery which stand outside particular theoretical or treatment paradigms enables such effects to be more readily explored.

(c) Group design

Although large group designs have an important function in comparing the effectiveness of different treatment approaches, they also have problems. The principal ones of concern here have to do with the extent to which one can generalise the findings of such studies to individuals. As Hersen & Barlow (1976) have pointed out, there are two problems here: first, one cannot automatically infer that the results from a large homogenous group of subjects are necessarily representative of the population of such subjects, and, second, it is difficult to take the average response of a group of subjects and generalise to the individual case. Averaged data can often result in misrepresentation of individual functions (Sidman, 1960), a point of considerable relevance to psychotherapy research where such an effect obscures improvement and deterioration in outcome within a group (Bergin, 1966).

The present study is, of course, subject to these difficulties in generalising findings to individuals. Single-case designs are frequently advanced as a useful alternative to group designs (e.g. Leitenberg, 1973; Hersen & Barlow, 1976) but, as far as the issue of generalisation to other individuals is concerned, they are also not immune from difficulties. They do nevertheless have the advantage of offering greater flexibility in investigations. As Hersen & Barlow (1976) point out:

"...the investigator can begin an immediate search for the cause of an experimental behaviour trend by altering his experimental designs on the spot. This feature...can provide immediate information on hypothesized sources of variability.' (p.40)

An important source of information that such designs can supply, especially when informed by time-series methods of analysis (Jones et al., 1977), has

to.../

has to do with the variability and fluctuations that are frequently observed within the treatment process, and which is not available to pre and post models. In the present study, it will be recalled, the observational data pointed to precisely such within-treatment variability.

Single-subject designs, however, are not being recommended here as a panacea. What is suggested is that group and single-subject approaches have their respective contributions in this area of enquiry. Kiesler (1981) has made a plea for such multidimensional research and has recommended a variety of research strategies as heuristic, valuable and necessary. He also makes the point that no single study is worth much in itself. Scientific progress results from empirical attacks on a problem. As Gelso (1979) suggests: 'Each study will be flawed, but we will come to 'know' the domain, albeit temporarily, through the cumulative efforts of experiments, hopefully each with a different set of methodological shortcomings.' (p. 13)

(d) Statistical significance

A principal concern in the foregoing section is the implications for individuals of data derived from groups of subjects. A similar issue arises in the context of discussions of statistically significant effects of treatment, where it may be argued that clinical significance for individuals is possibly lacking. In the present study, for example, it was suggested that even although task-attention data showed statistically significant improvements, these did not appear to have the (clinical) significance which would have altered teachers' perceptions accordingly.

Reliance on statistical inference has been the subject of much criticism, not just within the context of the clinical versus statistical significance debate (Cronbach, 1975), but also, for example, in critiques of the manner in which the emphasis on statistical significance in editorial policies creates serious publication biases (Bakan, 1967). Meehl (1978) recently suggested that this kind of over-reliance was one of the 'worst things that ever happened in the history of psychology'.

The observation in the present study of a gap between these forms of significance suggests a need in research of this nature to improve and expand the meanings of differences between treatments beyond those suggested by probability levels. One important method of determining clinical significance, which has been receiving increasing attention recently, is that of social validation (Kazdin, 1977b). This involves comparing the performance of the client with that of a 'normative' group of peers similar to the way that an individual's psychological test score may be compared with the 'norms' available for a sample of the general population. The question for social validation is whether an individual's improvement brings him within the range that would be deemed acceptable within his peer group. Comparison with normative groups could therefore increase the value of investigations such as the present one, although it should be pointed out that considerable care needs to be exercised in the choice of comparative group. Interventions with delinquents, for example, may call for a choice between the local norms of a sub-culture and those of society at large.

(e) Integrity of treatment

As already pointed out, a weakness of the present study is the absence of adequate data - observational or otherwise - which could have thrown light on the degree and quality of treatment implementation. Petersen et al. (1982) have pointed to a curious double standard that has developed in behavioural research where much attention is lavished on the operational definition and reliability of measurement of the dependent variable or target response, but the same rigour is lacking when the same behaviours appear as independent variables. These authors point to the inadequacy of statements that 'treatment was applied as outlined in the method section or treatment manual', and indicate several ways in which the quality of treatment implementation may deteriorate. Thus, a treatment agent may become lax with timing or with effortful techniques, or may begin to add procedures which have not been prescribed. This kind of problem may be especially relevant to studies.../

studies attempting to transfer laboratory-proved procedures to the field, where control over treatment implementation is likely to be reduced. Boruch & Gomez (1977) have suggested that the frequent failures of such attempted transfers may be due to decreased treatment integrity.

Provision of more adequate information in the present study would have required considerably more extensive coverage by the observer of teacher behaviour. The available manpower did not permit this, unfortunately. Petersen et al. (1982) suggest that some workers may feel that the potential cost of the inaccuracy involved in the omission of such information, must be balanced against the cost of ensuring accurate independent variable application. The cost-effectiveness of the solution would dictate whether it would be applied or not. These authors argue, however, that deliberately accepting inaccuracy in basic subject matter because it may be inconvenient to gather accurate information has no place in a science of behaviour. It is hard to disagree with this, and in the interests of increasing the meaningfulness of evaluation data, such controls need closer attention.

(v) Practical Recommendations

This section will examine some of the practical implications of the present study for teachers, psychologists, administrators and others within the educational context. These will be drawn from considerations in three areas: first, the findings on the evaluative measures; second, the wider issues in introducing behaviour modification into regular schools; third, the accommodation of criticisms and the revision of the present behavioural model.

(a) Findings on evaluative measures

As indicated earlier in this chapter there is little in the results of this study that can be taken as strong positive recommendations for the methods used. In general, this may be interpreted as reflecting the limitations of social reinforcement with children of secondary school age. This does not imply that token reinforcement is therefore necessary as a stronger and .../

and more powerful means of influencing pupil motivation. It is felt that social reinforcement is worthy of retention, if only for its contribution in lessening coercive approaches to control. Within the behavioural paradigm, however, it may need to be supplemented by a greater emphasis firstly on antecedents which influence behaviour and, secondly, on consequences other than 'social' or 'token' consequences which occur naturally within the school environment. Where antecedents are concerned, it is evident that behaviour is influenced by contextual events preceding it as well as by consequences (Wheldall & Austin, 1980). This may involve, for example, paying greater attention to the quality of the curriculum and ensuring that it generates sufficient interest, so that 'discipline' and 'control' are sought <u>through</u> pupils' activities rather than independently of it; examining activity management in terms of smoothness of delivery and transitions, as in Kounin's (1970) work; investigating the role of different seating arrangements and their influence on behaviour (e.g. Wheldall, 1981b). These are just a few possibilities which might be pursued. Antecedent factors were by no means ignored in the present study but it is possible that the intervention may have been strengthened had they been given greater consideration.

Pursuing factors which may serve as consequences for behaviour, it is evident that there are a number of possibilities within the school framework. Access to preferred activities is one important one but secondary schools' organisation is such that such 'rewards' are often difficult to mobilise because they are rigidly frozen into the timetable. Another possibility is making use of free time as a reinforcer but, again, many school authorities would view it with disfavour. The writer successfully employed free time in this fashion in a separate study, but it was felt that the procedure was viable only because the children in question spent most of their day in a special unit attached to the school, to which different 'rules' were applied. There are obvious difficulties in this kind of approach but it is quite possible that if free time serves as significant motivation for learning, then pupils could end up learning more in less time. 'Free' time is not necessarily wasted time.

Taking a closer look at the implications of results on specific measures, one area where social reinforcement could be seen to have made a useful contribution is in the sphere of children's problems of isolation. Adults' approval and encouragement may be of considerable help, therefore, to children who are cut off from peer interaction and activities. The impact of such help may be even greater in teaching settings which are more informal than the chalk-and-talk framework which dominated the present study.

On a less positive note, the present approach, with its focus mostly on children's behaviour, cannot be recommended where a prime target is improvement in academic achievement. A more direct focus on academic behaviour or on curricular matters may be necessary.

None of the changes on other measures seem sufficiently clear cut, nor are specific to behaviour modification, to enable firm recommendations based on them to be made.

(b) <u>Introducing behaviour modification into regular schools</u>

In general, what is recommended here is that one needs to be aware of the many other processes and influences in school life which co-exist along with behavioural programmes and, for better or worse, impinge upon it. Some of these have been given detailed consideration in an earlier section, so only a brief outline of recommendations will be offered here.

First, adequate training and consultation require time and this must be planned for at the outset. While not understating the need for personal commitment, arrangements which encroach on teachers' 'free' or personal time may ultimately create stresses which undermine the programme. Arrangements involving release from teaching, especially at the early training stage, probably offer the optimum conditions.

Second, in view of the numerous other potential influences on teacher behaviour in the classroom, it may be useful to develop some understanding of the teacher's own personal style, preferably from observation within the classroom. This can serve as an important corrective for the notion that.../

that behavioural procedures can be straightforwardly applied by any teacher in any lesson in any classroom. Such observation can give important insights into how intended management procedures need to be adapted to a particular teacher's circumstances and personal idiosyncracies. Observation in the classroom also affords useful opportunities in detecting the various classroom processes which may be contributing to the presenting complaint.

Third, in addition to the immediate influences operative in the classroom, attention needs to be paid to the wider influences which operate at the level of the colleague group or the social system of the school. Thus, colleagues' views may need to be taken into account, and recognition must be given to the possible pressures - subtle or otherwise - which the individual teacher may feel. Training and consultation in pairs or small groups may provide some support for teachers and help to establish a core of interested staff. However, this may not necessarily involve any direct handling of conflict within the wider colleague group. Informal contacts with other school staff may be helpful here, even at the level of staffroom discussion. It is difficult to see how more formal and organised discussions could be justified for small-scale interventions. It is apparent, however, that educational psychologists are increasingly pursuing system-wide interventions rather than limited involvements with individual referrals (e.g. Gillham, 1981; Figg & Ross, 1981). Such an approach is ideally suited to the kinds of issues raised here.

(c) <u>Accomodating criticisms and re-appraising the intervention model</u>

The practical recommendations offered so far have been derived from consideration of the results of the study and experience in attempting to implement the treatment model. It is necessary to supplement these with recommendations which take into account some of the criticisms levelled against the approach in terms of both its standpoint towards treatment, and also the model of inquiry which is implicit in it.

These criticisms have been considered in some detail in an earlier section..../

section. With a view to improving interventions of the present kind, or future research, what recommendations can be drawn from them? The main thread, which seems to run through most of the criticisms, is the need for a greater involvement of the 'client' - broadly defined - in determining the shape of the research and intervention process. This thrust is apparent whether one considers calls for improved status for 'internal' or 'mental' events, involvement of children and teachers in discussions of objectives, ensuring that their respective needs are considered, or recommendations of a drastically revised model of science and research enquiry. The same thrust is apparent in some of the alternative models of helping which were briefly considered earlier, such as Glasser's (1977) approach, Gordon's (1974) Teacher Effectiveness Training and Spivack & Shure's (1974) problem-solving methods.

The writer feels that greater client involvement - in its widest sense - would be to the advantage of future research within the classroom context, but it is less clear precisely how this should be pursued. It could be pursued within the behaviour modification tradition, or within the newer alternative framework as epitomised by the 'ethogenic' approach (Harre & Secord, 1972) or 'new paradigm' research (Reason & Rowan, 1981). It is not clear whether it could be pursued within an amalgam of both since their standpoints are so conflicting.

Within the behavioural perspective, developments are possible in several directions. First, work in the area of self-reinforcement (e.g. Douglas et al., 1976; Nelson & Birkimer, 1978) can be seen as offering individuals greater personal control over their behaviour, and, in lessening dependence on the availability of external reinforcement, it offers the potential - as yet unfulfilled - of facilitating generalisation across settings and maintenance in time. Second, the burgeoning area of cognitive behaviour modification (e.g. Kendall & Hollon, 1979) gives individual and idiosyncratic private events a legitimate place that was previously denied them. Third, problem-solving approaches (e.g. Spivack et al., 1976; Platt et al., 1974) emphasise.../

emphasise the client's own generation of alternative ways of dealing with conflict situations. Finally, contracting approaches offer greater equality in the organisation of helping plans. It should also be noted that some researchers, such as Graubard et al. (1971) have reversed the traditional paths of influence in the classroom by evolving procedures whereby pupils could influence and modify teacher behaviour. However, this is equally one-sided. The overall pattern of results in the present study would tend to suggest that, while more research is undoubtedly needed, it should not be just more of the same. The success of the present model of intervention is probably too limited and patchy to justify a call for independent replication, followed by wholesale application. So which is to be the preferred direction? In a situation where a familiar framework begins to feel constricting and limited, and an alternative one seems attractive but relatively untested, there are a number of options. Principally, these involve remaining with the safe and familiar framework, opting wholesale for the relatively unknown, or inching slowly towards the new and different way of working. In the present context, the latter seems the most preferable, in the same way that it is wiser for a non-swimmer to get accustomed to the shallow end before risking floundering in the deep.

From this perspective, the problem-solving model (e.g. Spivack & Shure, 1974) offers a useful and promising stepping stone between the two frameworks. It can be seen as offering a number of advantages. First, it retains the strengths of behaviour modification in its emphasis on task-analysis and step-wise approach to objectives (however these are determined). Second, since the teacher is centrally involved in mediating this kind of help, it has the same advantage of 'reach' as behavioural procedures. Third, in stressing an anticipatory approach to problems, it has important preventive implications. Its possible application with young children also reinforces its preventive potential. Fourth, the model implicitly involves children in discussion with teachers as the means of shaping the helping process. Fifth, it has the potential of establishing itself within the school curriculum and..../

and, as such, discussion of interpersonal relationships could become an integral part of classroom activities, significantly widening the concept of curriculum. The impact of such a model of intervention - and indeed of any model - is likely to be enhanced by a recognition of its co-existence alongside many other influences, philosophies and social pressures operating within school life. Such recognition has major implications for the manner in which a helping model is introduced, how staff training and support are conducted, and the form that the research process eventually takes.

(d) Concluding note

It is worth noting that the making of any recommendations from a study such as the present one needs to be tempered by the recognition that research knowledge is only one kind of evidence or understanding. Teachers and others who work within school settings will gather their own 'evidence' on the basis of their own learning, their experiences, and day-to-day happenings. The teachers who participated in this study would no doubt want to square their 'evidence' with any results or recommendations the writer might produce. Research evidence is not to be seen as superior to theirs, but as a different and additional source of understanding upon which they can draw.

(vi) Summary

The present chapter has offered interpretations of results based both on changes in individual measures and on overall trends, and on the relative performance of the treatment regimes. Some of the problems encountered in introducing behavioural procedures into the regular school setting were discussed, along with problems considered more widely in terms of criticisms directed at the behavioural approach and the model of scientific enquiry it espouses. Some methodological implications were then outlined, and, finally, a number of practical recommendations were offered.

CHAPTER 9

CONCLUSION

CONCLUSION

The behavioural programme investigated here was one form of intervention in a comparative treatment study conducted within ordinary secondary schools. The treatments were viewed as having the potential of reaching the large numbers of children within the schools suggested by a variety of surveys to be in need of special help. Alternatives were sought to the traditional approach of individual psychotherapy within child guidance which had not been shown to be conspicuously successful. Behaviour modification commended itself as one approach and has been the focus of concern in the present study, with group counselling and parent counselling/teacher consultation being viewed as treatment comparison regimes rather than being fully explored in their own right.

Behaviour modification seemed especially well-suited to the thrust of this study, in presenting a positive and constructive orientation towards children's problems and involving the teacher centrally as the mediator of procedures. This latter feature is very much in keeping with the recent trend to helping children with special needs within the ordinary school framework. It additionally reflects an attempt to capitalize on the wider socialising function of schools. This implies a widening of the teacher's role but the behavioural approach was viewed as eminently compatible with many routine teaching activities. It was recognised that behaviour modification was but one approach to children's problems, and, as such, it might not be consistent with individual teachers' own views on the tasks of management.

A particularly attractive feature of the behavioural approach was the growing body of findings which pointed to successes with the range of problems with which the study was concerned. However, little of this evidence had been gathered with children of secondary school age, in regular settings. This study set out to explore how a large-scale application of the model might fare in ordinary secondary schools, in circumstances which replicated as.../

as closely as possible the constraints and limitations of the natural environment as opposed to one given over to a major degree to research activities. Limited time for training and consultation with teachers was seen as one important component of the natural school setting. The treatment procedures heavily emphasised social reinforcement, which had been relatively unexplored with children of this age. An attempt was made to fill other gaps in previous research by employing multiple measures in evaluation and conducting long term follow-up.

These were the principal themes underlying the behavioural programme. How may its outcome be judged? A number of issues need to be considered here. First, the results for the present model are not particularly favourable, in that, while a number of advantages over untreated controls and parent-counselling/teacher consultation were recorded, many of them are shared with group counselling. They are not therefore specific to behaviour modification. The extent of shared success with group counselling is intriguing and indicates a surprising commonality in outcome for two therapies with avowedly discrepant theoretical bases. Research into children's own understanding and perceptions of therapy might yield valuable insights here. Second, the consistent superiority of behaviour modification over parent-counselling/teacher consultation (within its range of 'successes') suggests that the specifically behavioural components within the consulltation process may have proved a critically important part of the regime. Third, amongst the findings on individual measures, it was particularly noteworthy that behaviour modification had greater impact than all other regimes in helping isolated children in the short term. However, it made no significant impact in the sphere of reading comprehension (other than relative to PC). Fourth, in addition to the lack of independent success, there is no evidence that changes were maintained over a period of time, although some advantages were observed even at the last follow-up. Interpretation of advantages emerging at this late stage, however, is difficult, and to attribute them solely...../

solely to treatment variables is to ignore other potentially important factors. Along with the similar outcome for two different treatments in this study, such findings suggest that investigations of behaviour change need to go beyond the possibly limited explanations offered by specific theories underlying particular treatment orientations. Thus change may occur as a function of other factors intervening in time, outwith conventionally recognised 'treatment' frameworks, or as a function of processes that are common to different treatment approaches. Theoretical preconceptions may at times lead to the neglect of such factors.

There is therefore no compelling evidence that the present time-limited training and consultation model had a substantial and enduring independent impact on the problems confronted in this study. It is quite possible that social reinforcement procedures may be of less value with secondary school children than with younger children in primary school settings. Secondary pupils may find peer reinforcement much more attractive than anything the teacher has to offer. The dearth of published studies employing social reinforcement with older children may reflect this relative lack of effectiveness.

On a wider note, pertaining to the methodology of the study, two points are of interest. First, the finding of significant differences on <u>all</u> the outcome measures is ample justification for the inclusion of a variety of measures in assessing change. Change can be multi-dimensional and much information may be lost in studies which limit assessments and restrict their focus. Second, the emergence of positive outcomes in the long term (despite difficulties in interpretation) shows the importance of conducting follow-up assessments. Behavioural treatments just like their counterparts under the psychotherapy umbrella (cf. Wright et al., 1976), may set processes in motion the results of which are not apparent until well beyond the period of intervention.

Data yielded by outcome measures, however, are but one source of 'evidence' in evaluating interventions. As indicated in the Discussion section, the writer's experience in conducting the research led to a number of insights into the constraints which operate within the complex social setting of the regular school framework. The straightforward operations of the laboratory are hard to reproduce in natural settings. The niceties of treatment procedures and implementation may have the foremost priority in the researcher's mind but, in practice, many other processes require attention. In the school setting, teachers' personal and social realities call for consideration and research priorities need to be viewed within the context of teachers' tasks, responsibilities and relationships. Lieberman (1980) suggests that:

'After almost twenty years of behavioural analysis and therapy, workers in the field must realize that political, personal and social factors determine upwards of 90% of the success and survival of technical procedures.....Implementation, survival and dissemination of empirically validated interventions require much more than data and journal publications.' (pp 370-371).

Such considerations, along with the outcome data, suggest to the writer that interventions within regular schools require considerable breadth of awareness and sensitivity in those initiating change. Such a requirement has many facets, but, perhaps, principally, it suggests a need to get to know and understand teachers' working environments very thoroughly. A different thrust for research is also encouraged by criticisms of behaviour modification and the model of enquiry it represents. These principally suggest a need to involve participants in research (teachers, other school staff, and children in the present context) in a dialogue about research processes, goals and objectives.

There are perhaps more questions to be answered at the end of the present study than there were at the beginning. This may be a reflection of the quality of the study. On the other hand, it may also reflect something of the reality of research which may answer some questions along the way, but, in so doing, points up new issues, perspectives and complexities the more the subject is explored. New knowledge seems to be accompanied inevitably by fresh awareness of areas of ignorance. So the research.../

research cycle is pursued, hopefully with a more informed and balanced outlook, and an increased tolerance for uncertainty.

APPENDICES

APPENDIX 1: SCREEN MEASURES

Further information will be provided here about the screen measures which will amplify that provided in the text. In addition, copies of the measures are appended.

(i) Sociometric measures

The sociometric measure was presented to children in the following form:

NAME.....

YOUR CLASS

Here are some questions about your class. Read them carefully.

These questions are about children you like to do things with. You will like to do things with some children but not with others. But different children like to do things with children you do not choose.

Fill in every space with the name of a child in your class. You can choose any one you wish, even someone who is away from school today. You can put down the same name for more than one question if you would really choose the same person. Remember, do not leave any blank spaces and choose only children from YOUR class.

1 (a) Which 3 children would you most like to SIT BESIDE in your class?

1..... 2..... 3.....

(b) Which 3 children would you not like to SIT BESIDE in your class?

1.....2..... 3.....

2 (a) Which 3 children from your class would you most like to PLAY WITH at playtime?

1..... 2..... 3.....

(b) Which 3 children would you not like to PLAY WITH at playtime?

1..... 2.....3.....

After the screening stage, large class charts, on which the children's names were listed and numbered in alphabetical order, were pinned to the blackboard so that choices could be indicated by/

by numbers rather than full names. This shortened administration time and obviated the need to check that children's full names had been written.

Children's popularity scores were calculated by summing choices received for 1a and 2a. Any children receiving 0 or 1 choice were classified as 'isolated'. Rejection scores were calculated by summing choices received for 1b and 2b. Children receiving 12 or more choices were classified as 'rejected'. Test-retest reliability was examined by conducting assessments, one month apart, with a total of 61 children in two classes of 11-12 year olds. Correlations of 0.72 for Isolation and 0.87 for Rejection were found.

Some data on the relationships between these two sociometric indices and the other screen measures in the present study were calculated for a sample of 625 children, comprising one-third of the screen positives, and one-third of screen negatives (Macmillan et al., 1980). For Isolation, correlations of 0.21, 0.19 and 0.15 were observed with Rutter Total, Rutter Neurotic and Rutter Antisocial respectively. The corresponding correlations for Rejection were 0.28, 0.23 and 0.26. These correlations, though small, were all highly significant ($p<0.001$). No significant correlations, however, were observed with JEPI Neuroticism. The pattern of correlations observed with the teacher ratings suggests that isolation may be less closely related than rejection to the kind of disturbed behaviour which teachers identify. Lack of popularity also may not be synonymous with poor adjustment, and the lower test-retest reliability of Isolation suggests that it may be generally less stable and enduring than Rejection.

(ii)<u>Teacher rating: the Rutter B2 scale Rutter (1967)</u>

This scale has already been described in the text. It yields a Total score, and two subscale scores, for Neurotic and Anti-social behaviour. To increase the reliability of the two subscales, the number of items contributing to them was enlarged in the present study (Kolvin et al., 1981). This extension followed an inspection of Rutter's (1967) data, bearing on discrimination of the two types of disorder. Items were added on the basis that they (a) differentiated the psychiatric group from controls and (b) differentiated the diagnostic groups. The revised subscales.../

subscales consisted of the following items for the neurotic subscale: G, H, J, K,N, Q, R, V and W; and for the antisocial behaviour: A, B, D, E, O, P, S, T and Z.

For the B version of this scale, which differs from B2 only in the wording of a few items, Rutter (1967) reported test-retest reliability for total scores over a two-month period as 0.89. He also refers to the work of Richman (1964) who reported a thirteen-week test-retest reliability of 0.85 using a slightly modified version of the scale. Rutter also reported an inter-rater correlation of 0.72 with a two to three month gap. Richman reported inter-rater reliability of 0.70.

(iii) Personality measure: The Junior Eysenck Personality Inventory (Eysenck, 1965)

This measure has already been described in the text. The dimension of concern in the present study is neuroticism. Split-half reliabilities for neuroticism with the 11-12 year old group are in the region of 0.84, and lie between this figure and 0.82 for age groups up to 15 years. Test re-test reliabilities average between 0.7 and 0.8. These were obtained on 1,056 boys and 1,074 girls in all, the time elapsing between test and re-test being one month.

Principal component analyses were conducted by Eysenck for children of different ages and it was concluded that neuroticism could be measured adequately at all the ages covered by the inventory. Validating data are not presented in the manual, although it is reported that in an assessment of 229 child guidance cases, this group was very significantly above the standardization group with respect to neuroticism.

In the correlations conducted amongst all screen criteria, JEPI Neuroticism achieved significant relationships with none of the other criteria. How might this be interpreted? It is possible that JEPI Neuroticism does not validly identify disturbed children. It may, however, identify children who are indeed disturbed, but whose difficulties are not tapped by the other screen criteria. Clinical ratings on children selected solely by the JEPI criterion support this latter interpretation (Macmillan et. al., 1980).

RUTTER S-C-A-L-E B (2)

To be completed by Teachers

Name of Child ..... Boy/Girl School .....

Date of Birth ..... Form .....

Below are a series of descriptions of behaviour often shown by children. After each statement are three columns:- "Doesn't Apply", "Applies Somewhat" and "Certainly Applies". If the child definitely shows the behaviour described by the statement place a cross in the box under Column 2 "Certainly Applies". If the child shows the behaviour described by the statement but to a lesser degree or less often place a cross in the box under Column 1 "Applies Somewhat". If, AS FAR AS YOU ARE AWARE, the child does not show the behaviour, place a cross in the box under Column 0 "Doesn't Apply".

Please complete on basis of child's behaviour IN THE PAST 6 MONTHS

Put ONE cross against EACH statement. Thank you.

	0 Doesn't Apply	1 Applies Somewhat	2 Certainly Applies
A Very restless, has difficulty staying seated for long	☐	☐	☐
B Truants from school	☐	☐	☐
C Squirmy, fidgety child	☐	☐	☐
D Often destroys or damages own or others' property	☐	☐	☐
E Frequently fights OR is extremely quarrelsome with other children	☐	☐	☐
F Not much liked by other children	☐	☐	☐
G Often worried, worries about many things	☐	☐	☐
H Tends to be on own - rather solitary	☐	☐	☐
I Irritable. Touchy. Is quick to 'fly off the handle'	☐	☐	☐
J Often appears miserable, unhappy, tearful or distressed	☐	☐	☐
K Has twitches, mannerisms, or tics of the face or body	☐	☐	☐
L Frequently sucks thumb or finger	☐	☐	☐
M Frequently bites nails or fingers	☐	☐	☐

	STATEMENT	0 Doesn't Apply	1 Applies Somewhat	2 Certainl Applies
N	Tends to be absent ~~from school for trivial~~ reasons	☐	☐	☐
O	Is often disobedient	☐	☐	☐
P	Cannot settle to anything for more than a few moments	☐	☐	☐
Q	Tends to be fearful or afraid of new things or new situations	☐	☐	☐
R	Fussy or over-particular child	☐	☐	☐
S	Often tells lies	☐	☐	☐
T	Has stolen things on one or more occasions in the past 12 months	☐	☐	☐
U	Unresponsive, inert or apathetic	☐	☐	☐
V	Often complains of aches or pains	☐	☐	☐
W	Has had tears on arrival at school OR has refused to come into the building in the past 12 months	☐	☐	☐
X	Has a stutter or stammer	☐	☐	☐
Y	Resentful OR aggressive when corrected	☐	☐	☐
Z	Bullies other children	☐	☐	☐

Is there anything else unusual about the child's behaviour? - or are there any other comments you would like to make?

.....

.....

.....

.....

Signature: Mr/Mrs/Miss .....

JUNIOR EYSENCK PERSONALITY INVENTORY

by Sybil B. G. Eysenck

NAME .....

AGE..... SEX.....

E= ☐ N= ☐ L= ☐

Instructions

Here are some questions about the way you behave, feel and act. After each question is a space for answering "YES" or "NO".

Try to decide whether "YES" or "NO" is your usual way of acting or feeling. Then put a cross in the circle under the column headed "YES" or "NO". Work quickly, and don't spend too much time over any question. Be sure not to leave out any questions.



UNIVERSITY OF LONDON PRESS LTD
Warwick Lane, London E.C.4

E O N O L O

REMEMBER TO ANSWER EACH QUESTION

	YES	NO
1. Do you like plenty of excitement going on around you?	O	O
2. Do you often need kind friends to cheer you up?	O	O
3. Do you nearly always have a quick answer when people talk to you?	O	O
4. Do you sometimes get cross?	O	O
5. Are you moody?	O	O
6. Would you rather be alone instead of meeting other children?	O	O
7. Do ideas run through your head so that you cannot sleep?	O	O
8. Do you always do as you are told at once?	O	O
9. Do you like practical jokes?	O	O
10. Do you ever feel "just miserable" for no good reason?	O	O
11. Are you rather lively?	O	O
12. Have you ever broken any rules at school?	O	O
13. Do lots of things annoy you?	O	O
14. Do you like doing things where you have to act quickly?	O	O
15. Do you worry about awful things that might happen?	O	O
16. Can you always keep every secret?	O	O
17. Can you get a party going?	O	O
18. Do you get thumping in your heart?	O	O
19. When you make new friends do you usually make the first move?	O	O
20. Have you ever told a lie?	O	O
21. Are you easily hurt when people find fault with you or the work you do?	O	O
22. Do you like telling jokes or funny stories to your friends?	O	O
23. Do you often feel tired for no good reason?	O	O
24. Do you always finish your homework before you play?	O	O
25. Are you usually happy and cheerful?	O	O
26. Are you touchy about some things?	O	O
27. Do you like mixing with other children?	O	O
28. Do you say your prayers every night?	O	O
29. Do you have "dizzy turns"?	O	O

	YES	NO
30. Do you like playing pranks on others?	○	○
31. Do you often feel fed-up?	○	○
32. Do you sometimes boast a little?	○	○
33. Are you mostly quiet when you are with others?	○	○
34. Do you sometimes get so restless that you cannot sit in a chair long?	○	○
35. Do you often make up your mind to do things suddenly?	○	○
36. Are you always quiet in class, even when the teacher is out of the room?	○	○
37. Do you have many frightening dreams?	○	○
38. Can you usually let yourself go and enjoy yourself at a gay party?	○	○
39. Are your feelings rather easily hurt?	○	○
40. Have you ever said anything bad or nasty about anyone?	○	○
41. Would you call yourself happy-go-lucky?	○	○
42. Do you worry for a long while if you feel you have made a fool of yourself?	○	○
43. Do you often like a rough and tumble game?	○	○
44. Do you always eat everything you are given at meals?	○	○
45. Do you find it very hard to take no for an answer?	○	○
46. Do you like going out a lot?	○	○
47. Do you sometimes feel life is just not worth living?	○	○
48. Have you ever been cheeky to your parents?	○	○
49. Do other people think of you as being very lively?	○	○
50. Does your mind often wander off when you are doing a job?	○	○
51. Would you rather sit and watch than play at parties?	○	○
52. Do you find it hard to get to sleep at nights because you are worrying about things?	○	○
53. Do you usually feel fairly sure you can do the things you have to?	○	○
54. Do you often feel lonely?	○	○
55. Are you shy of speaking first when you meet new people?	○	○
56. Do you often make up your mind when it is too late?	○	○
57. When children shout at you, do you shout back?	○	○
58. Do you sometimes feel specially cheerful and at other times sad without any-good reason?	○	○
59. Do you find it hard to really enjoy yourself at a lively party?	○	○
60. Do you often get into trouble because you do things without thinking first?	○	○

PLEASE CHECK TO SEE THAT YOU HAVE ANSWERED ALL THE QUESTIONS

APPENDIX 2: ESTABLISHING CUT-OFFS AND WEIGHTINGS FOR SCREEN CRITERIA

It was decided to take extreme scores on each of the screen measures as indicators of deviance. The actual scores used as cut-offs on the Rutter scale and the JEPI were decided by examining published data on the characteristics of the instrument. For the subscales of the Rutter questionnaire and sociometry, however, cut-off scores were decided on the basis of a pilot study of 200 cases.

With the Rutter B2 scale, a cut-off of 9 has been regarded as providing the best discrimination between children attending child guidance clinics and a normal sample (Rutter et. al., 1970). This cut-off was retained and it was assigned a weighting of point 1 towards the deviance classification. In addition, the more extreme score of 15 was given a weighting of 2. This was arrived at by adding 1 standard deviation to Rutter's original cut-off.

Cut-offs were assigned on the neurotic and antisocial subscales which gave rise to a yield closest to that of the total score. This proved to be a score of 4 on each scale and this was assigned a weighting of 1. It should be emphasized that the use of the sub-scale as a weighting score is different from Rutter et al's use of the subscale for diagnostic purposes. The three Rutter weighted scores were added together to contribute to the total screen score.

With the sociometric criteria, decisions about cut-offs were guided by findings with the Rutter scale. Cut-offs were adjusted so that percentages similar to that identified by the Rutter Total cut-off were selected. For isolation, this proved to be a cut-off of one positive choice or less. For rejection, 12 or more negative/

negative choices was taken as the cut-off. Scores on or beyond these cut-offs were each weighted 1 point.

Cut-offs of 1½ and 2 standard deviations above the mean for neuroticism were taken for the JEPI, with the scores being rounded up to be equivalent for both sexes. These scores were 20 and 23 respectively. Children with scores of 20-22 were allotted 2 points, and those with more extreme scores of 23 or 24, 3 points.

APPENDIX 3: SECOND LEVEL MEASURES

(1) Devereux Elementary School Behaviour Rating Scale (D.E.S.B.)

Spivack and Swift (1967) designed the D.E.S.B. to measure overt behaviours which reflect a child's overall adaptation to the demands of the classroom setting and which may affect his achievement in that setting. Since the intention of the authors was to measure behaviours of specific relevance to the classroom and not behaviours relating to disturbance in a psychiatric context, the initial pool of 111 items for study was derived from a series of discussions with 72 normal and special class teachers.

Two initial factor analyses of behaviour ratings, one on data from normal and the other on data from special classes for emotionally handicapped children were completed on 579 children (Spivack and Swift, 1966). The same factors emerged for both settings, and were replicated in a further factor analysis with 1325 normal French children (Swift et. al., 1972).

The D.E.S.B. consists of 44 items, defining 11 behavioural factors, the items in each factor being selected on the basis of factor loadings, similar patterns of correlation with other variables such as I.Q. and sex, and significant correlation with academic achievement. A copy of the scale is appended. Fig. A3-1 shows the factor composition and the items which contribute to each factor.

The teacher is asked to make each behaviour rating based on his own norm of what the average, normal child of the same sex and age behaves like in the classroom. The scale takes about 8 minutes to complete. Norms are based on ratings by 32 teachers in grades from kindergarten through to sixth grade (i.e. ages 5-11 years) for/

for 809 public school children. Test-retest reliabilities over a one-week period are reported as ranging between .85 and .91 for all D.E.S.B. factors with a median test-retest coefficient of .87. Inter-rater reliability, based upon 40 pairs of ratings in a normal classroom, ranged from .62 to .77 with a median rater reliability of .70.

As regards validity, Spivack and Swift report that each of the 11 factors has been shown to correlate significantly with teacher grades after the influence of IQ has been partialled out, with normal American and French children, and with groups of emotionally disturbed children.

FIGURE A3(1): Devereux Elementary School Behaviour Rating Scale:

Factor Composition and Content

Behavior Factor	Factor Item Raw Scores	Tot'l Raw Sc.
1. Classroom Disturbance	needs control 11 ___ 13 ___ interfere teases 12 ___ 30 ___ drawn in	
2. Impatience	starts 1 ___ 44 ___ go back sloppy 36 ___ 47 ___ rushes	
3. Disrespect-Defiance	disrespect 5 ___ 9 ___ subject defy t'ch'r. 7 ___ 16 ___ rules	
4. External Blame	t'ch'r. help 2 ___ 34 ___ blames called on 25 ___ 38 ___ too hard	
5. Achievement' Anxiety	test scores 22 ___ 31 ___ testing right answ. 23 ___ 33 ___ sensitive	
6. External Reliance	see others 24 ___ 42 ___ swayed rely t'ch'r. 29 ___ directions 32 ___ 46 ___ choices	
7. Comprehension	understands 10 ___ 37 ___ recites applies 35 ___	
8. Inattentive - Withdrawn	lose attn. 18 ___ 28 ___ oblivious not attnd. 20 ___ 43 ___ reachable	
9. Irrelevant - Responsiveness	exagg. story 14 ___ 17 ___ interrupt answers 15 ___ 26 ___ irrel. talk	
10. Creative Initiative	brings in 3 ___ 6 ___ start disc. act. imag. 4 ___ 21 ___ talk exper.	
11. Need Closeness to Teacher	seeks t'ch'r. 8 ___ 39 ___ friendly helps 19 ___ 45 ___ phys. close	

FIGURE A3(1) (CONTINUED)

YOU ARE GOING TO RATE THE OVERT BEHAVIOR OF A STUDENT. FOR ITEMS 1-26 USE THE RATING SCALE BELOW. WRITE YOUR RATING (NUMBER) FOR EACH ITEM IN THE BOX TO THE LEFT OF THE ITEM NUMBER.

Very frequently	Often	Occasionally	Rarely	Never
5	4	3	2	1

COMPARED WITH THE AVERAGE CHILD IN THE NORMAL CLASSROOM SITUATION, HOW OFTEN DOES THE CHILD...

Rating	Item
☐	1. Start working on something before getting the directions straight?
☐	2. Say that the teacher doesn't help him enough (i.e., won't show him how to do things, or answer his questions)?
☐	3. Bring things to class that relate to current topic (e.g., exhibits, collections, articles, etc.)?
☐	4. Tell stories or describe things in an interesting and colorful fashion (e.g., has an active imagination, etc.)?
☐	5. Speak disrespectfully to teacher (e.g., call teacher names, treat teacher as an equal, etc.)?
☐	6. Initiate classroom discussion?
☐	7. Act defiant (i.e., will not do what he is asked to do, says: "I won't do it")?
☐	8. Seek out the teacher before or after class to talk about school or personal matters?
☐	9. Belittle or make derogatory remarks about the subject being taught (e.g., "spelling is stupid")?
☐	10. Get the point of what he reads or hears in class?
☐	11. Have to be reprimanded or controlled by the teacher because of his behavior in class?
☐	12. Poke, torment, or tease classmates?
☐	13. Annoy or interfere with the work of his peers in class?
☐	14. Tell stories which are exaggerated and untruthful?
☐	15. Give an answer that has nothing to do with a question being asked?
☐	16. Break classroom rules (e.g., throw things, mark up desk or books, etc.)?
☐	17. Interrupt when the teacher is talking?
☐	18. Quickly lose attention when teacher explains something to him (e.g., becomes fidgety, looks away, etc.)?
☐	19. Offer to do things for the teacher (e.g., erase the board, empty the pencil sharpener, open the door, get the mail, etc.)?
☐	20. Makes you doubt whether he is paying attention to what you are doing or saying (e.g., looks elsewhere, has blank stare or faraway look, etc.)?
☐	21. Introduce into class discussion personal experiences or things he has heard which relate to what is going on in class?
☐	22. Get openly disturbed about scores on a test (e.g., may cry, get emotionally upset, etc.)?
☐	23. Show worry or get anxious about knowing the "right" answers?
☐	24. Look to see how others are doing something before he does it (e.g., when teacher gives a direction, etc.)?
☐	25. Complain teacher never calls on him (e.g., that teacher calls on others first, etc.)?
☐	26. Make irrelevant remarks during a classroom discussion?

FIGURE A3(1) (CONTINUED)

FOR ITEMS 27-47 USE THE RATING SCALE BELOW:

Extremely	Distinctly	Quite a bit	Moderately	A little	Very slightly	Not at all
7	6	5	4	3	2	1

COMPARED WITH THE AVERAGE CHILD IN THE NORMAL CLASSROOM SITUATION, TO WHAT DEGREE IS THE CHILD...

Rating	Item
☐	27. Unable to change from one task to another when asked to do so (e.g., has difficulty beginning a new task, may get upset or disorganized, etc.)?
☐	28. Oblivious to what is going on in class (i.e., not "with it," seems to be in own "private" closed world)?
☐	29. Reliant upon the teacher for directions and to be told how to do things or proceed in class?
☐	30. Quickly drawn into the talking or noise-making of others (i.e., stops work to listen or join in)?
☐	31. Outwardly nervous when a test is given?
☐	32. Unable to follow directions given in class (i.e., need precise directions before he can proceed successfully)?
☐	33. Sensitive to criticism or correction about his school work (e.g., gets angry, sulks, seems "defeated", etc.)?
☐	34. Prone to blame the teacher, the test, or external circumstances when things don't go well?
☐	35. Able to apply what he has learned to a new situation?
☐	36. Sloppy in his work (e.g., his products are dirty or marked up, wrinkled, etc.)?
☐	37. Likely to know the material when called upon to recite in class?
☐	38. Quick to say work assigned is too hard (e.g., "you expect too much," "I can't get it," etc.)?
☐	39. Responsive or friendly in his relationship with the teacher in class (vs. being cool, detached or distant)?
☐	40. Likely to quit or give up when something is difficult or demands more than usual effort?
☐	41. Slow to complete his work (i.e., has to be prodded, takes excessive time)?
☐	42. Swayed by the opinion of his peers?
☐	43. Difficult to reach (e.g., seems preoccupied with his own thoughts, may have to call him by name to bring him out of himself)?
☐	44. Unwilling to go back over his work?

COMPARED WITH THE AVERAGE CHILD IN THE NORMAL CLASSROOM SITUATION, TO WHAT DEGREE DOES THE CHILD...

Rating	Item
☐	45. Like to be close to the teacher (e.g., hug or touch the teacher, sit or stand next to teacher, etc.)?
☐	46. Have difficulty deciding what to do when given a choice between two or more things?
☐	47. Rush through his work and therefore make unnecessary mistakes?

(ii) Barker Lunn Children's Attitude Scales - S7

These attitude scales were developed in the context of a study by the National Foundation for Educational Research into the effects of streaming and non-streaming in junior schools (Barker Lunn, 1967).

The questionnaire contains ten attitude scales: attitude to school; interest in school work; importance of doing well; attitude to class; 'other' image of class; conforming versus non-conforming pupil; relationship with teacher; anxiety in the classroom; social adjustment; academic self-image. Each scale is made up of 6-10 statements made by children during group discussions and selected by factor analysis and scalogram analysis. The items contributing to each scale are shown in the appended reproduction of the questionnaire. (The questions are grouped here to illustrate the composition of the scales, so this is not the order in which the questions are presented. 'Keyed' responses indicate which response categories are scored.)

Some 2,300 third and fourth year junior school children in 28 schools completed the final form of questionnaire. Inter-correlation of the scales shows two clusters; the one dealing with attitudes towards school and school work and the other with social relationships and the personality of the pupil. The internal consistency of the scales was determined by Cronbach's Alpha-coefficient, yielding a range of 0.69 to 0.90 with a mean of 0.81.

A large number of correlations are reported with other measures, such as teacher and parent ratings of ability; sociometric data, achievement scores and interest scores. The expectation of a relationship between school performance and attitude is borne out. All the scales correlate significantly with test scores in English, Problem Arithmetic (all except conforming/non-conforming), Essays, Verbal Reasoning and Non-Verbal Reasoning. A similar pattern is

evident for teacher ratings, and also, with the exception of conforming/non-conforming, with parent ratings. The 'social adjustment' scale correlated 0.21 with sociometric status. Interest scores correlated most highly with 'attitude to school' and 'interest in school work'.

Composition of the Barker-Lunn children's Attitude Scales

Scale 1: Attitude to school

Item No.	Statement	Keyed Response
41	School is fun	Always
25	School is boring	Hardly ever
56	I like school	Yes
18	I bet going out to work is better than school	No/Not sure
60	I would leave school tomorrow if I could	No/Not sure
48	Going to school is a waste of time	No/Not sure

Scale 2: Interest in school work

Item No.	Statement	Keyed Response
12	We spend too much time doing arithmetic	Hardly ever
23	I like doing hard sums	Yes, often
64	At school they make you do things you don't want to do	No
31	We have interesting lessons in school	Most of the time
47	I enjoy most school work	Yes
8	School lessons are boring	Never/Sometimes
47	I enjoy most school work	Yes/Not sure

Scale 3: Importance of doing well

Item No.	Statement	Keyed Response
43	I should like to be one of the cleverest pupils in the class	Yes = 2, Not sure = 1
44	I work and try very hard in school	Always = 2, Most of time = 1
52	I should like to be very good at school work	Yes = 2, Not sure = 1
63	Doing well at school is most important to me	Yes = 2, Not sure = 1
11	I should like to be better at games than at school work	No = 2, Not sure = 1

Scale 4: Attitude to class

Item No.	Statement	Keyed Response
59	I like being in my class	Yes = 2, Not sure = 1
40	I'd prefer to be in another class	No = 2, Not sure=1
34	I hate being in the class I'm in now	No = 2, Not sure = 1
26	I'm happy to be in the class I'm in now	Yes = 2, Not sure = 1
16	I think a lot of children of my age would like to be in my class	Yes = 2, Not sure = 1
9	My class is nicest of all	Yes = 2, Not sure = 1
13	I'd rather be in my class than the others for my age	Yes = 2, Not sure = 1
19	I shall be sorry to leave my class	Yes = 2, Not sure = 1

Scale 5: 'Other' image of class

Item No.	Statement	Keyed Response
29	Other children think we're very clever in my class	Yes
53	Other children make fun of my class	Never
62	Other classes think they're better than us	No/Not sure
55	When people ask me what class I'm in I always feel happy to tell them	Yes
37	My class gets blamed for things we don't do	False/Not sure
15	Other classes think we're nice in my class	Yes/Not sure

Scale 6: Conforming versus non-conforming pupil

Item No.	Statement	Keyed Response
6	When the teacher goes out of the room I play about	Never
33	I dislike children who are noisy in class	Yes
3	It's nice to fool about in class	Never
35	I like children who get into trouble	No
22	I like people who get me into mischief	No
35	I like children who get into trouble	No/Not sure

Scale 7: Relationship with teacher

Item No.	Statement	Keyed Response
36	Teacher is interested in me	Yes
4	Teacher gets on well with me	Most of the time
51	Teacher thinks I'm a trouble maker	No
54	I think my teacher likes me	Yes
49	My teacher is nice to me	Most of the time
24	Teacher is always nagging me	No
4	Teacher gets on well with me	Most of the time/ Sometimes
51	Teacher thinks I'm a trouble maker	No/Not sure

Scale 8: Anxiety in the classroom situation

Item No.	Statement	Keyed Response
32.	Children who can't do their school work feel ashamed	No
58	It would bother me if I got my work wrong	No
27	School work worries me	No
38	I should feel a little afraid if I got my spellings or sums wrong	No
61	I enjoy being asked questions by my teacher	Yes/Not sure
38	I should feel a little afraid if I got my spellings or sums wrong	No/Not sure
27	School work worries me	No/Not sure
20	I'm scared to ask my teacher for help when I don't understand	Never/Sometimes

Scale 9: Social adjustment

Item No	Statement	Keyed Response
46	I don't always get on well with some of the children in my class	False
39	I think the other children in my class like me	Yes
10	I have no one to play with at playtime	Never true
21	I have no friends that I like very much in my class	False

Scale 10: Academic self-image

Item No.	Statement	Keyed Response
5	I get a lot of sums wrong	Hardly ever = 2, Sometimes = 1
7	I think I'm pretty good at school work	Yes = 2, Not sure = 1
50	I'm useless at school work	Never = 2, Sometimes = 1
17	My teacher thinks I'm clever	Yes = 2, Not sure = 1
45	I'm very good at sums	Always = 2, Sometimes = 1
30	When we have tests I get very good marks	Most of the time = 2, Sometimes = 1
42	I find a lot of school work difficult to understand	Hardly ever = 2, Sometimes = 1
14	I sometimes think I'm no good at anything	False = 2, Not sure = 1
57	I don't seem to be able to do anything really well in school	False = 2, Not sure = 1

(iii) General Ability Test (National Foundation for Educational Research

This is an 80-item ability test, with alternating verbal and non-verbal items (40 of each), yielding a verbal, a non-verbal and a total score. It is designed for group administration, with 40 minutes' completion time. For both this test and the Reading Comprehension Test, there are no published norms. This is a drawback for descriptive purposes, but in the present study they were used solely as comparative change measures. Two sample pages are appended.

At age eleven, test-retest reliability of 0.94 was reported by Douglas (1967), with a sample of 363 children. Barker-Lunn (1970) reported test-retest reliability of 0.93 with 253 nine-year-olds, and 0.95 with 254 ten-year-olds.

With regard to validity, Douglas (1967) reported a correlation of 0.93 with the N.F.E.R. Verbal Test 8A (11+ selection test) with a group of 74 children. In addition, he reported from survey data with an eleven-year-old sample of 3,418 children, correlations of 0.69 with both a Mechanical Reading Test and a Vocabulary Test, and 0.75 with an Arithmetic Test. (Details of these tests are not given.)

SAMPLE ITEMS FROM THE GENERAL ABILITY TEST

1

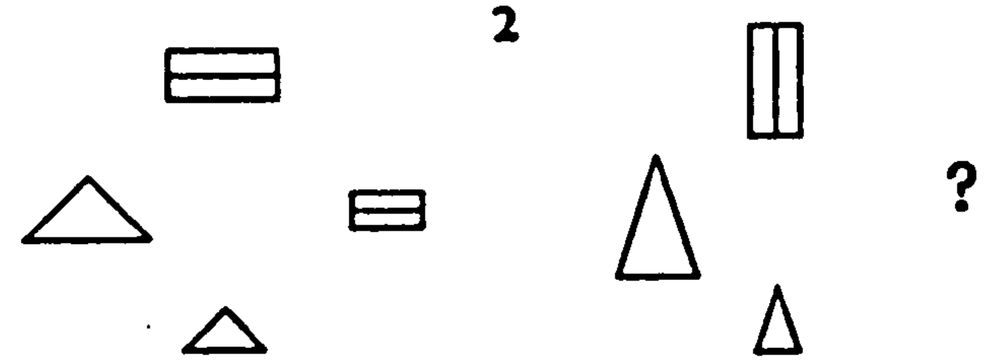
Mouse — Cat

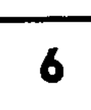
Dog — Squeak — Sheep — ?

Bark — Bleat

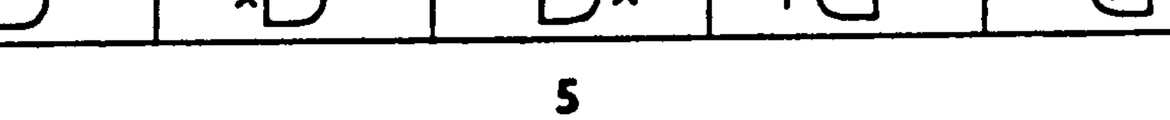
Roar	Goat	Cow	Mew	Grunt

2



3



4

Penny — Cauliflower

Farthing — Coin — Cabbage — ?

Shilling — Turnip

Carrot	Sprout	Garden	Crops	Vegetable

5

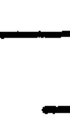
May — June

March — July — April — ?

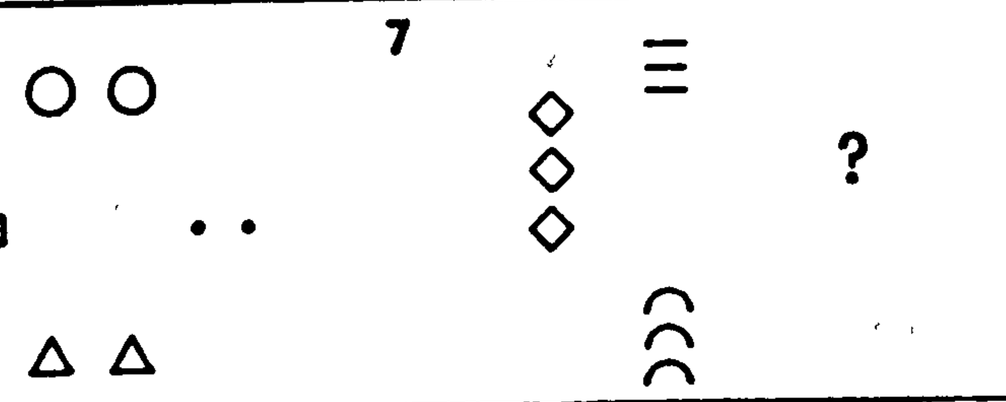
January — February

Summer	August	July	September	Month

6

7



8

Rink — Leaves

Ink — Irate — Eaves — ?

Rate — Late

Elate	Eat	Weaves	Veal	Sale

GO STRAIGHT ON TO THE NEXT PAGE

73

Snout

Nose

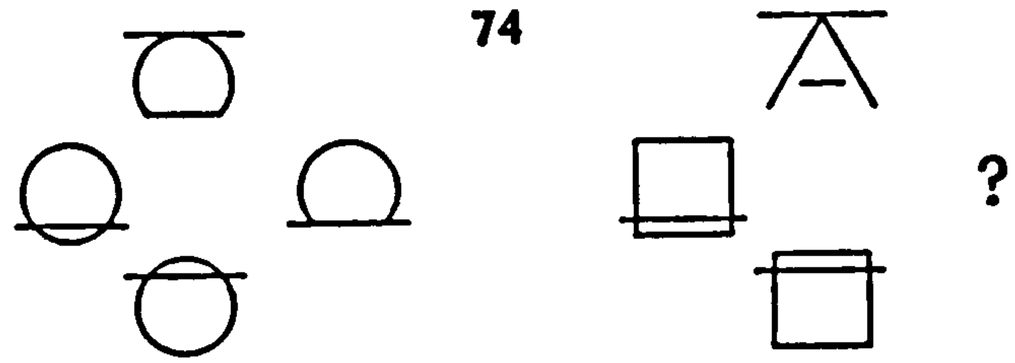
Pig Trotters Man ?

Hide

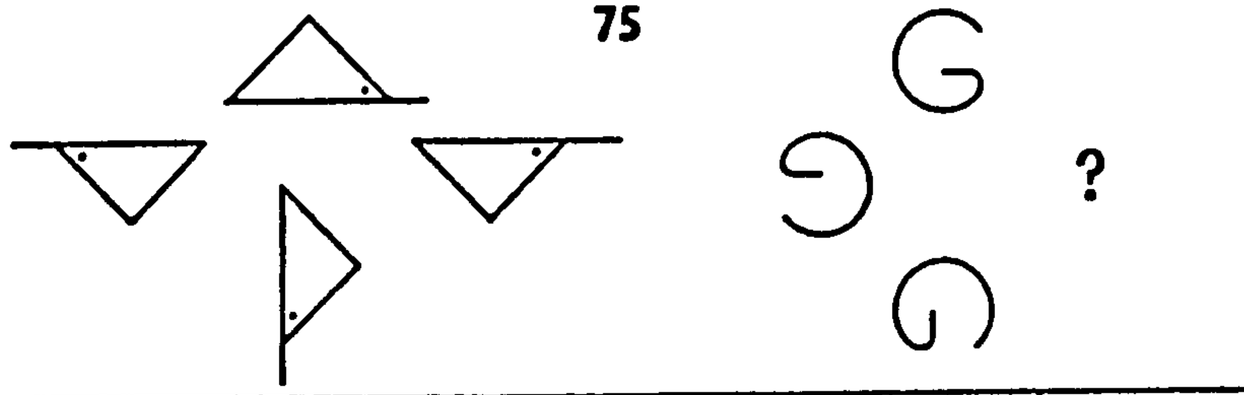
Skin

Paws	Knees	Feet	Trippers	Trousers

74



75



76

Trial

Judge

Detection Prison Detective ?

Crime

Criminal

Warder	Capture	Police	Escape	Reporter

77

Train

Link

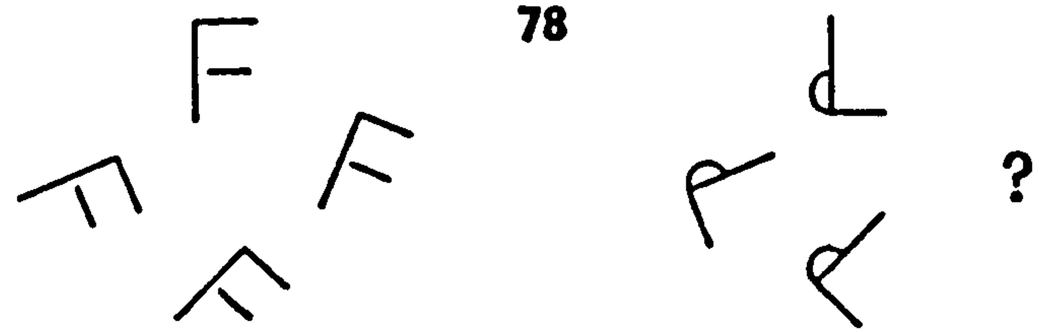
Rain Neat Ink ?

Eat

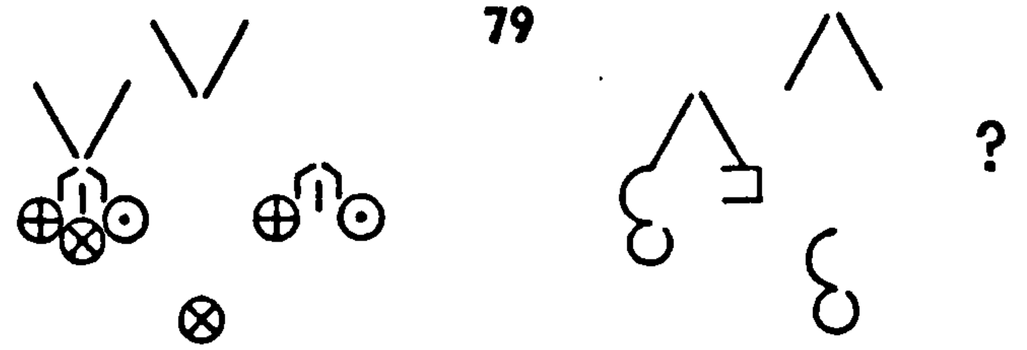
Eel

Line	Lean	Leek	Kneel	Keel

78



79



80

Put

Nut

Tap Tip Tan ?

Pot

Not

Ten	Tin	Tun	Net	Nat

END OF TEST

(iv) Reading Comprehension Test (National Foundation for Educational Research

This 35-item test of reading comprehension was constructed as a parallel test to the Watts-Vernon (Start and Wells, 1972) so that for each item in the Watts-Vernon, there was an item of comparable facility value in the parallel test. It is designed for group administration and 15 minutes' administration time is allowed. Two sample pages are attached.

A test-retest correlation of 0.90 was reported by Douglas (1967) for a sample of 124 eight-year-olds. In a survey conducted by Douglas, some 3,418 children were tested at eleven years. For this group, the following correlations were reported: 0.56 with Picture Test 1 (N.F.E.R.) - a 60-item non-verbal picture test; 0.87 with a Mechanical Reading Test; and 0.68 with a Vocabulary Test.

411

SAMPLE ITEMS FROM THE READING COMPREHENSION TEST

Do not write here.

8. When Jane learnt that she had passed the examination she felt very (sorry, considerate, interesting, pleased, pleasant). 8.....

9. She had been dieting for a month, but her weight had not (shown, increased, shrunk, decreased, grown). 9.....

10. The peaceful atmosphere of the village was (dismembered, disregarded, disputed, disrupted, dissected) when the huge lorry rumbled down the street. 10.....

11. The painter was not very successful financially although it was (necessary, obvious, scrupulous, reasonable, disappointing) that he was very talented. 11.....

12. Despite the (convenience, expense, poverty, economy, pleasure) of living in this area, I have managed to save some money. 12.....

13. All the (individuals, inhabitants, decisions, thoughts, dwellers) of the village were cut off by the blizzard. 13.....

14. After some time had elapsed, we once again (renewed, helped, lapsed, involved, deserted) our efforts to solve the difficult problem. 14.....

15. The building which was going to be demolished had been (derisive, animated, vicious, derelict, derivative) for some months. 15.....

16. Considering the great capacity that a human being has for love, it is astounding that the history of mankind should be frequently blackened by deeds of (love, manliness, consideration, hate, history). 16.....

17. The man said that he was not interested and did not want to (partition, participate, purport, produce, predict) in the discussion. 17.....

18. The (process, proceeds, pages, gain, progress) from the sale of the book were given to a worthwhile charity. 18.....

TURN OVER WITHOUT WAITING TO BE TOLD

Do not write here.

29. When the speaker asked if there were any questions he was (involved, immured, inundated, implied, instructed) with queries. 29.....

30. To identify any form of government with a religion is a dangerous error, for it confounds the permanent with the transitory, the (good, temporary, real, absolute, religious) with the contingent. 30.....

31. Having had her expectations so much raised, it was very (realistic, discrediting, uplifting, disconcerting, discriminating) to have them suddenly shattered. 31.....

32. A strong gale had been blowing all night and the (following, ensuing, much, causing, destroying) damage in the garden took John several days to repair. 32.....

33. Whilst self-restraint is a praiseworthy quality which one should try to cultivate, it runs as much danger of being exaggerated as its (counterpart, equal, quality, virtue, denial), the free expression of personal feelings and emotions. 33.....

34. It took three men to (refrain, prohibit, restrain, catch, restrict) the youth from rushing into the burning shed to rescue his pet dog. 34.....

35. The boss firmly denied any accusation of discrimination; he claimed that the policies of his firm were not dictated by (shareholders, prejudice, incrimination, accusations, profits). 35.....

END OF TEST

APPENDIX 4 : OBSERVATIONAL PROCEDURES

(i) Behavioural categories. The categories employed in the coding of children's and teachers' behaviour are shown in Tables A4(1) and A4(2) respectively. The children's categories and definitions were drawn from the work of Becker et al. (1967). The one substantial modification to their system was the inclusion of the category of 'unresponsive' behaviour to cover situations where children failed to respond to specific requests. Teachers' categories were drawn from Madsen et al. (1968), not in as complete a fashion as the children's categories. 'Other response' was included to cover behaviours not classifiable as approval or disapproval. Also, 'no response' seemed a necessary addition.

(ii) Method of observation. The observer had a clipboard, ratings sheets and a stop watch. An illustration of the rating sheet is given in Fig. A4(1). Child categories were precoded down the left hand side of the page. The recording procedure required the observer to observe a child for ten seconds and then in the next five seconds record both the child's behaviour and the teacher's response to it (in terms of the categories for teacher behaviour). Thus each cell represented ten seconds' observation and conveyed information about the behaviour of both teacher and child. It was hoped that this manner of recording would illustrate the nature of typical interactions between them. Having recorded one child's behaviour in this way, the observer would then move on to the next child for the following ten seconds, then the next, and so on. Having covered all the target children, the cycle would begin again, and continue until the end of the lesson.

(ii) Computing inter-observer agreement. Agreement between observers was computed with the formula:

$$\frac{\text{number of agreements}}{\text{number of agreements + number of disagreements}} \times 100$$

This was computed for task-relevant behaviour, which was the dependent variable of prime concern in this part of the study. In addition, a computation was made for the level of agreement that could be expected by chance. If the base rate of the behaviour in question is high, then the opportunity of yielding agreements by chance alone is inflated. Chance agreement was calculated by squaring the base rate of the two categories of task-relevant behaviour and the aggregated off-task behaviours and summing these values (Johnson & Bolstad, 1973) i.e. R^2 + non-R^2, where R is task-relevant behaviour.

These computations were conducted for each child (N=33) involved in the observer-agreement exercise and are reported in Table A4(3). For teacher behaviour, the two categories of concern were 'approval' and 'disapproval'. Agreements between the observers for each of the ten teachers observed are shown in Table A4(4). It will be noted that 'not applicable' appears opposite several of the teachers. This indicates that no instances of the behaviour in question were observed in the recording session, so agreement could not be computed. This does not mean that these teachers were not employing approval or disapproval, only that these were not directed at the target children <u>at the point of observation</u> during the session. Chance agreement rates are not quoted here because the incidence of both praise and disapproval was so low that the chance agreement rate is quite negligible.

As pointed out by Hartmann (1977), there is no entirely agreed upon set of rules for deciding upon an acceptable level of observer agreement. He indicates that 80% agreement has some consensus among behavioural researchers, while Jones et al. (1975) have suggested that 70% agreement is adequate where complex codes are used. The mean agreement achieved for children in the present exercise was 92%, well above these suggestions of minimum level, and well above the mean chance agreement level of 64%.

x^2 comparisons were conducted to determine whether the observed degree of agreement was better than chance. (The test focused on the number of observation intervals in which agreement was found as against the number in which chance agreement was expected). The outcome showed that only three of the trials failed to better chance levels. See Table A4(3).

(iv) Validity As a means of exploring the validity of the observational procedures, correlations were computed between baseline levels of task-relevant behaviour and baseline teacher ratings on the two Devereux dimensions of Classroom Disturbance and Inattentive-Withdrawn for the sample of 55 children. Product-moment correlations of -.05 and .12 respectively, were obtained, neither of which reached statistical significance.

TABLE A4(1): DIRECT OBSERVATION: PUPIL CATEGORIES

Symbol	Category	Definitions
X	Gross motor behaviour	Getting out of seat; standing up, moving around out of chair; rocking in chair; disruptive movement without noise.
N	Disruptive noise with objects	Tapping pencil or other objects; clapping, stamping feet; rattling or tearing paper; banging books on desk. (Rate only if noise can be heard with eyes closed. Do not include accidental noise, or noise made while performing X above.)
A	Disturbing others directly and aggression	Grabbing others' work/materials; knocking neighbour's book off desk; destroying others' property; kicking; hitting; shoving; pinching; slapping;

TABLE A4(1) (Continued)

Symbol	Category	Definition
		slapping; striking or poking with object; throwing object at another person; pulling hair; tripping.
L	Looking around, distraction. (non task)	Turning head or head and body more than 90° to look at another person, or the back of the room; showing objects to another child; attending to another child; looking into space. (Exclude when turning or attending to another child occurs when under teacher's instructions, or in academic context.)
V	Blurting out, commenting and	Answering teacher without raising hand or without being called on; making comments or remarks when no question has been asked; calling teacher's name to get attention; crying, screaming, singing; whistling; laughing loudly; negative comments towards teacher. (Must not be directed to another child but may be directed to teacher.)
T	Talking	Carrying on conversations with other children when it is not permitted. (Must be directed to a particular child or children.)
D	Unresponsive	No response to teacher when asked questions, or to make a contribution to the lesson.
R	Task relevant behaviour	On task, e.g. writing answers, answering questions, listening, raising hand; other behaviours clearly permitted by the teacher in the carrying out of an academic task.

TABLE A4(2): DIRECT OBSERVATION: TEACHER RESPONSE CATEGORIES

Symbol	Category	Definition and examples
2	Approval	(a) Contact. Positive physical contact such as patting, holding arm or hand.
		(b) Praise. Verbal comments indicating approval, commendation or achievement, e.g. that's good, you're doing fine, you are studying well.
		(c) Facial attention. Smiling at a child.
		(d) Feedback. Giving feedback for academic correctness.
3	Disapproval	(a) Holding the child. Forcibly holding the child, putting him outside, grabbing, hitting, slapping, shaking.
		(b) Criticism. Critical comments of high or low intensity, yelling, scolding, raising voice, e.g. Don't do that, stop talking, quiet!
		(c) Threats. Consequences mentioned by the teacher to be used at a later time. "If _____ then _____" comments.
		(d) Facial attention. Frowning, scowling or grimacing at a child.
0	Other response -	Calling on a child for an answer. Probing for a response. Non-critical instruction: one which does not imply disapproval.
1	No response -	The teacher makes no response to the behaviour observed, classifiable as one of above.

TABLE A4(3) : OBSERVER AGREEMENT FOR TASK RELEVANT BEHAVIOUR

Child	Percentage Agreement	Chance Agreement	Number of Observation Intervals	x^2 value	p <
1	100	61	60	14.96	.01
2	97	52	60	23.37	.01
3	100	52	60	26.58	.01
4	98	56	60	18.90	.01
5	80	64	64	2.56	n.s.
6	98	96	64	0.11	n.s.
7	100	68	64	9.64	.01
8	100	65	64	12.06	.01
9	100	78	64	3.97	.05
10	88	62	64	6.98	.01
11	96	70	60	17.14	.01
12	100	56	60	11.06	.01
13	81	53	60	20.74	.01
14	100	74	48	4.38	.05
15	100	74	48	4.38	.05
16	98	72	48	4.50	.05
17	98	74	48	3.73	n.s.
18	81	68	60	10.58	.01
19	82	62	60	5.96	.05
20	91	84	60	5.68	.05
21	84	67	60	4.00	.05
22	100	70	60	7.71	.01
23	100	74	60	5.48	.05
24	96	76	60	5.79	.05
25	80	55	60	6.82	.01
26	96	50	60	25.39	.01
27	100	54	60	23.51	.01
28	86	53	60	12.83	.01
29	100	50	60	30.00	.01
30	100	70	48	6.17	.05
31	100	74	48	4.38	.05
32	100	68	48	7.22	.01
33	83	56	48	6.24	.05
	Mean percentage agreement = 92.00				
	Mean chance agreement = 63.75				

Note: (1) One observation interval equals ten seconds.

(2) x^2 was computed for the difference between the number of observation intervals in which agreement was recorded and the number of intervals in which agreement could be expected by chance.

FIG. A4(1) Illustration of the rating sheet employed in direct observation

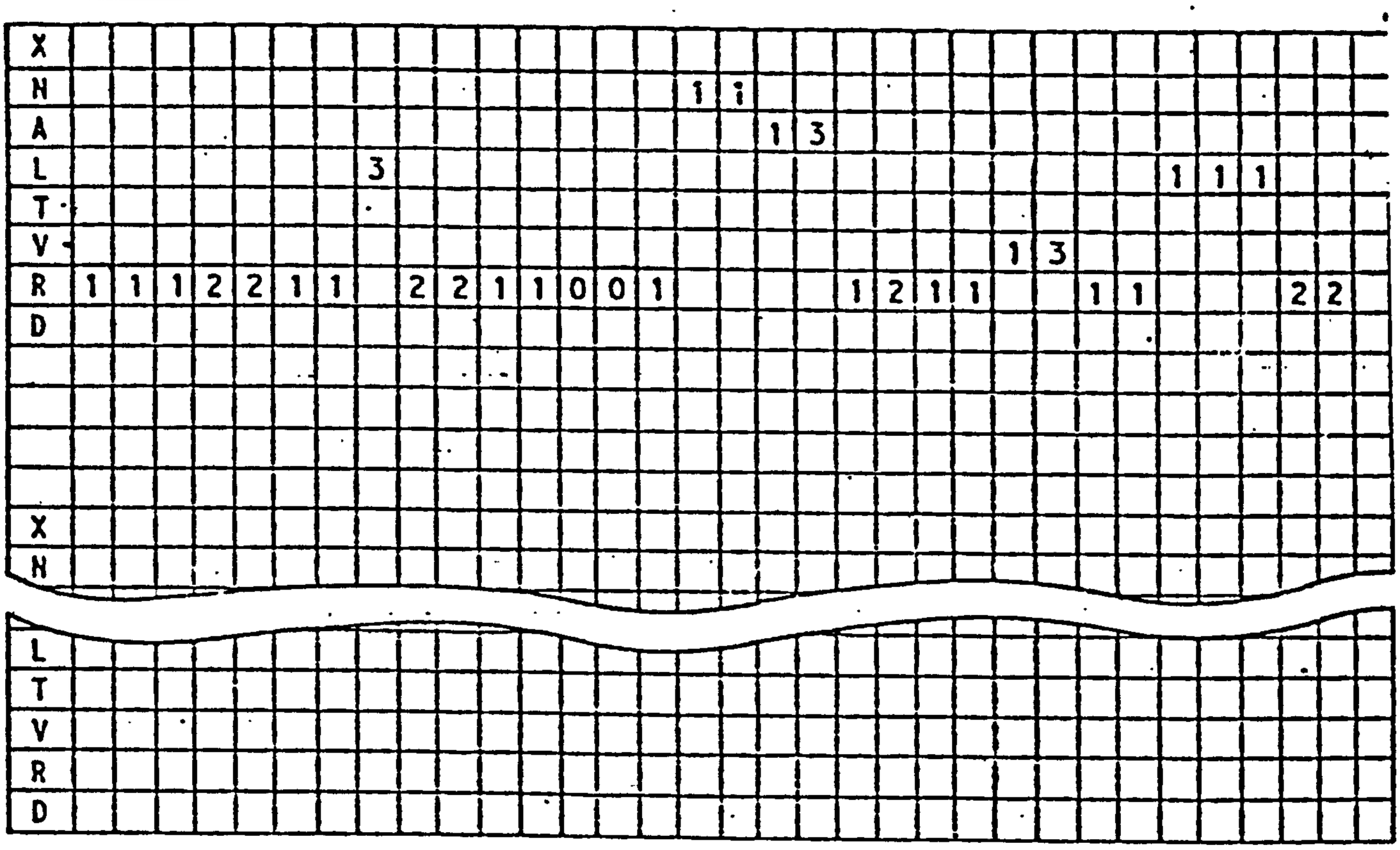


TABLE A4(4) : OBSERVER AGREEMENT FOR TEACHER APPROVAL AND DISAPPROVAL

	Percentage Agreement	
Teacher	Approval	Disapproval
1	100	n/a
2	100	n/a
3	100	80
4	n/a	83
5	n/a	n/a
6	100	100
7	n/a	82
8	100	100
9	89	100
10	100	n/a
Mean percentage agreement	98%	91%

Note: n/a means 'not applicable'. No instances of the behaviour were observed in the session.

APPENDIX 5: SUMMARY OF CONTENT OF TRAINING DOCUMENT AND TRAINING SEMINARS

(i) Document content.

Introduction: putting BM in context of range of classroom management procedures; basic assumptions; range of application.

Characteristics of BM: (a) emphasis on working with behaviour (contrast with psychodynamic models); problems in defining behaviour

(b) functional analysis: putting behaviour in an environmental context - eliciting factors and consequences for behaviour.

(c) intervention in real life settings: the role of natural caretakers; contrast with office-bound therapies.

(d) emphasis on measurement and evaluation: the role of observational data; methods of observation; experimental designs.

Principles of behaviour modification: Basis in: operant conditioning; classical conditioning; modelling and imitation learning.

Methods of strengthening behaviour: (a) positive reinforcement; types of reinforcement; considerations in application, (b) shaping, (c) chaining, (d) negative reinforcement, (e) contracting procedures, (f) modelling.

Methods of weakening behaviour: (a) Extinction - problems in application; praise and ignore approaches, (b) punishment - limits and side-effects; time-out procedures; response cost.

Practical problems in applying BM in the classroom: (a) Making social reinforcement work effectively - consistency, contingency, finding 'rewards' that work; teachers' attitudes to social reinforcement, (b) feasibility of token reinforcement programmes in the classroom/

classroom, (c) problems of generalisation - across classes, into the home; maintenance.

(ii) Seminar content.

The training document was circulated to all teachers before the seminars began, with the request that it should have been read before the first seminar. Three seminars, each of about an hour's duration, were held. Much of the seminar content consisted of highlighting specific parts of the document, with illustrations from relevant literature. Since participation and discussion were encouraged, the emphases and extent of coverage on specific points varied a little from group to group. The following is a brief outline of the seminar content.

Seminar 1. Basic assumptions and characteristics of B.M., justifying the model; the teacher's central role in management of problem behaviour. Defining problem behaviour; setting objectives; observing behaviour; functional analysis.

Seminar 2. Principal methods for (a) strengthening behaviour and (b) weakening behaviour; praise and ignore procedures; the trap of disapproval; soft reprimands. (Examples from the literature.)

Seminar 3. Applying behavioural procedures in the regular classroom; practical problems in application; examples of possible approaches to selected cases presented by the teachers.

APPENDIX 6: TESTING FOR HOMOGENEITY OF WITHIN-GROUP RECESSION

Before conducting analyses of covariance to test hypotheses 1-6, preliminary analyses concerning the assumption of homogeneity of regression were necessary. The initial score proved to be the most important covariate, in having most effect on improvement. Homogeneity of regression with this covariate was therefore examined first to determine which measures could be validly analyzed by analysis of covariance.

The Rutter B2 Total showed heterogeneous regressions at midline and final follow-ups. However, the regimes did not differ significantly at midline by either analysis of variance or analysis of covariance. At the final follow-up, covariance produced more conservative results than analysis of variance, so the covariance results were adopted.

The Rutter B2 antisocial measure showed heterogeneous regressions only at the midline follow-up, at which point the regimes did not differ significantly by either covariance or variance analysis. The regimes differed significantly at final follow-up, and as the variance and covariance analyses were consistent, the latter were adopted.

The JEPI Neuroticism measure showed heterogeneous regressions at all three follow-up points. Similar significant results on analyses of variance and covariance were found at end of treatment and midline follow-ups. No significant differences were observed at final follow-up by either method of analysis, so the covariance findings were accepted.

The/

The three measures of verbal, non-verbal and total ability all showed heterogeneity of regression at the midline follow-up. At this follow-up, however, analysis of covariance, for all three measures, produced more conservative results than analysis of variance so its findings were accepted.

None of the present change measures was affected by inspection of homogeneity of regression upon the other covariates - general severity, non-verbal ability score, and social functioning within the family.

APPENDIX 7 : ANALYSIS OF OBSERVATIONAL DATA ON TEACHER BEHAVIOUR

Changes in teacher behaviour were not included in the main experimental hypotheses of this study. The observational procedure presented a difficulty for analysis in that observation points for teachers referred to their behaviour at that time only in relation to the target child under observation rather than towards any child in the classroom. This restricted focus, in retrospect, may have militated against the demonstration of changes in teacher behaviour.

Analyses of variance for repeated measures were carried out separately for rates of approval, disapproval, and for approval/disapproval ratios, for the baseline phase and the first, second and third phases of treatment, as with the children's behaviour. Since the amount of observational data gathered for individual teachers tended to vary, mean rates and ratios were calculated separately for schools rather than for individuals. The analyses of variance focused on school means. The approval/disapproval ratio was calculated as a way of inspecting whether the thrust of the intervention towards increasing approval rates, and decreasing disapproval rates, was realised. Overall means for rates of approval across baseline and first, second and third phases of treatment were as follows: 3.22%, 2.16%, 2.09% and 2.64%. The analysis of variance proved non-significant ($F = 1.18$, df 3/16).

For disapproval, the respective means were: 2.06%, 0.95%, 1.25% and 0.95% The analysis of variance again produced a non-significant result ($F = 1.82$, df 3/16). The respective means for/...

for approval/disapproval ratios were: 1.56, 2.27, 1.67 and 2.78. Analysis of variance again produced a non-significant result F ($F = 0.57$, df 3/16).

It is evident, therefore, that neither approval nor disapproval rates, nor their ratios, altered significantly during the intervention, at least as reflected in the data collection procedure adopted. As pointed out earlier, a wider focus on teacher behaviour may have facilitated the detection of any changes which were occurring. It is not possible, then, to relate the changes noted in children's task-related behaviour to variations in teacher behaviour.

REFERENCES

REFERENCES

Abidin, R.R. (1971) What's wrong with behaviour modification? J. Sch. Psychol. 9, 38-42.

Abramowitz, C.V. (1976) The effectiveness of group psychotherapy with children. Arch. Gen. Psychiat. 33, 320-326.

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: a factor analytic study. Psychol. Monogr., 80, 1-37.

Adams, R. & Biddle, B (1970). Realities of Teaching: Explorations with Video Tape. New York: Holt, Rinehart and Winston.

Allen, K.E., Benning, P.M. and Drummond, T. (1972) Integration of normal and handicapped children in a behavior modification preschool: a case study. Paper presented at the Third Annual Conference on Behavior Analysis in Education. Lawrence, Kansas.

Allen, K.E., Hart, B.M., Buell, J.S., Harris, F.R. and Wolf, M.M.(1964) Effects of social reinforcement on isolate behavior of a nursery school child. Child Development, 35, 511-518.

Allen, R.P., Safer, D.J., Heaton, R., Ward, A. and Barrell, M. (1975) Behavior therapy for socially ineffective children. J. American Academy of Child Psychiatry, 14, 500-509.

Anderson, R.J. (1974). School social work: the promise of a team model. Child Welfare, 53, 524-530.

Andrews, J.K. (1970). The results of a pilot program to train teachers in the classroom application of behavior modification technique. J. Sch. Psychol., 8, 37-42.

Arlin, M and Roth, G. (1978). Pupils' use of time while reading comics and books. Amer. Educ. Res. J., 15, 201-216

Aronfreed, J. (1976). Moral development from the standpoint of a general psychological theory. In T. Lickona (Ed): Moral Development and Behavior. New York: Holt, Rinehart and Winston.

Atkeson, B.M. and Forehand, R. (1978). Parents as behavior change agents with school related problems. Education and Urban Society, 10, 521-538.

Ayllon, T., Layman, D.and Burke S. (1972). Disruptive behavior and reinforcement of academic performance. The Psychological Record, 22, 315-323.

Ayllon, T., Layman, D. and Kandel, H.J. (1975). A behavioral-educational alternative to drug control of hyperactive children. J. Appl. Beh. Anal., 8, 137-146.

Ayllon, T., and Roberts, M.D. (1974). Eliminating discipline problems by strengthening academic performance. J. Appl. Beh. Anal. 7, 71-76.

Ayllon, T. and Rosenbaum, M.S. (1974). The behavioral treatment of disruption and hyperactivity in school settings. In: B.B. Lahey and A. E. Kazdin (Eds); Advances in Clinical Child Psychology, Vol. 1. New York: Plenum

Azrin, N. H. (1977). A strategy for applied research: learning based but outcome oriented. American Psychologist, 32, 140-150.

Azrin, N.H. and Holz, W.C. (1966). Punishment. In: W. K. Honig (Ed): Operant Behavior: Areas of Research and Application. New York: Appleton Century Crofts.

Baer, D.M., Wolf, M.M. and Risley, T.R. (1968). Some current dimensions of applied behavior analysis. J. Appl. Beh. Analysis, 1, 91-97.

Bakan, D. (1967). On Method: Toward a Reconstruction of Psychological Investigation. San Francisco: Jossey-Bass.

Bandura, A. (1974). Behavior theory and the models of man. American Psychologist, 28, 859 -869.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. Psychological Review, 84, 191-215.

Bandura, A., Grusec, J.E. and Menlove, F.L. (1967). Vicarious extinction of avoidance behavior. J. Pers. Soc. Psychol., 5, 16-23.

Barcai, A. and Robinson, E.H. (1969). Conventional group therapy with preadolescent children. Int. J. Group. Psychother., 19, 334-345.

Barcai, A., Vinbarger, C., Pierce, T. and Chamberlain, P. (1973). A comparison of three group approaches to underachieving children. Amer. J. Orthopsychiat. 43, 133-141.

Barker-Lunn, J.C. (Undated). Manual of instructions for use of primary children's attitude scales. Slough: N.F.E.R.

Barker-Lunn, J.C. (1967). Effects of streaming and non-streaming in junior schools. New Research in Education, 1.

Barker-Lunn, J.C. (1969). Development of scales to measure junior school children's attitudes. Brit. J. Educ. Psychol. 39, 64-71.

Barker-Lunn, J.C. (1970). Streaming in the Primary School. Slough: N.F.E.R.

Baron, A., Kaufman, A. and Stauber, K.A. (1969). Effects of instructions and reinforcement-feedback on human operant behavior maintained by fixed-interval reinforcement. J. Experimental Anal. Beh., 12, 701-712.

Barrett, C.L., Hampe, E. and Miller, L. (1978). Research on psychotherapy with children. In: S. L. Garfield and A. E. Bergin (Eds): Handbook of Psychotherapy and Behavior Change: An Empirical Analysis. (2nd ed.) New York: Wiley, pp 411-436.

Bauer, N.J. (1971). Differences in personality traits among most preferred and least preferred students in grades 10, 11 and 12. J. Educ. Res., 65, 65-70.

Baumeister, A.A. (1969). More ado about operant conditioning - or nothing. Mental Retardation, 7, 49-51.

Becker, W.C., Madsen, C.H. Jr., Arnold, C.R. and Thomas, D.R. (1967). The contingent use of teacher attention and praise in reducing classroom problems. Journal of Special Education, 1, 287-307.

Begelman, D.A. (1976). Behavioral classification. In: M. Hersen and A.S. Bellack (Eds): Behavioral Assessment: A Practical Handbook. New York: Pergamon, pp 23-50.

Berg, I., Consterdine, M., Hullin, R., McGuire, R and Tyrer, S. (1978) A randomly controlled trial of two court procedures in truancy. Brit. J. Criminol., 18, 232-244.

Berger, M. (1976). Learning Theories. In: M. Rutter and L. Hersov (Eds): Child Psychiatry: Modern Approaches. Oxford: Blackwell, pp 219-238.

Berger, M. (1979). Behaviour modification and professional practice: the dangers of a mindless technology. Bull. Brit. Psychol. Soc., 32, 418-419.

Bergin, A.E. (1966). Some implications of psychotherapy research for therapeutic practice. J. Abnorm. Psychol., 71, 235-246.

Bergin, A.E. and Lambert, M.J. (1978). The evaluation of therapeutic outcomes. In: S. L. Garfield and A. E. Bergin (Eds): Handbook of Psychotherapy and Behavior Change: An Empirical Analysis (2nd ed.) New York: Wiley pp 139-190.

Berleman, W.C., Seaberg, J.R. and Steinburn, T.W. (1972). The delinquency prevention experiment of the Seattle Atlantic Sheet Center. Social Service Review, 46, 323-346.

Berleman, W.C. and Steinburn, T.W. (1967). The execution and evaluation of a delinquency prevention program. Social Problems, 14, 413-423.

Best, K.G. (1973). Children's situational responses to the Junior Eysenck Personality Inventory. Unpublished M.A. Thesis. Univ. of Lancaster.

Bevins, S.M. (1970). A comparison of the effectiveness of individual and group counselling in the improvement of social adjustment of 5th and 6th grade children. Diss. Abs. Int. 30A, 5277

Biasco, F. (1966). The effects of individual counseling, multiple counseling and teacher guidance on sociometric status of children. Diss. Abs. Int. 27A, 323.

Bijou, S.W. and Baer, D.M. (1978). Behavior Analysis of Child Development. Englewood Cliffs: Prentice-Hall.

Birnbrauer, J., Wolf, M., Kidder, J. and Tague, C. (1965). Classroom behavior of retarded pupils with token reinforcement. J. Experimental Child Psychology, 2, 119-135.

Bjerstedt, A. (1956). The methodology of preferential sociometry. Sociometry Monograph No. 37.

Blackman, D. (1981). The experimental analysis of behaviour and its relevance to applied psychology. In: G. Davey (Ed.): Applications of Conditioning Theory. London: Methuen, pp 1-28.

Blanchard, E.B. and Johnson, R.A. (1973). Generalization of operant classroom control procedures. Behavior Therapy, 4, 219-229.

Boakes, R.A. and Halliday, M.S. (1970). The Skinnerian analysis of behaviour. In: R. Borger and F. Cioffi (Eds.): Explanation in the Behavioural Sciences. Cambridge: Cambridge Univ. Press, pp 345-374.

Bolstad, O.D. and Johnson, S.M. (1972). Self-regulation in the modification of disruptive classroom behavior. J. Appl. Beh. Anal. 5, 443-454.

Boruch, R. F. and Gomez, H. (1977). Sensitivity, bias and theory in impact evaluations. Professional Psychology, 8, 411-434.

Bowers, K.S. (1973). Situationism in psychology: an analysis and a critique. Psychol. Rev., 80, 307-336.

Bowles, P.E. and Nelson, R.O. (1976). Training teachers as mediators: efficacy of a workshop versus the bug-in-the-ear technique. J. School Psychology, 14, 15-26.

Box, G.E.P. (1953). Non-normality and tests on variance. Biometrika, 40, 318-335.

Brewer, W. (1974). There is no convincing evidence for operant or classical conditioning in adult humans. In: W. Weimer and D. Palermo (Eds.): Cognition and the Symbolic Processes. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Breyer, N.L. and Allen, G.J. (1975). Effects of implementing a token economy on teacher attending behavior. J. Appl. Beh. Anal. 8, 373-380.

Breyer, N.L., Calchera, D.J. and Cann,C. (1971). Behavioral consulting from a distance. Psychology in the Schools, 8, 172-176.

Brinich, P.M. (1981). Defining the units in observational research. Bull. Brit. Psychol. Soc., 34, 418-419

Broden, M., Hall, R.V., Dunlap, A., Clark, R. (1970). Effects of teacher attention and a token reinforcement system in a junior high school special education class. Exceptional Children, 36, 341-349.

Bronfenbrenner, U. (1943). A constant frame of reference for sociometric research. Sociometry, 6, 363-377.

Brophy, J. and Good, T. (1970). Teachers' communication of differential expectations for children's classroom performance: some behavioural data. J. Educ, Psychol. 61, 365-374.

Brophy, J.E. and Putnam, J.G. (1979). Classroom management in the elementary grades. In: D.L. Duke (Ed.): Classroom Management: The 78th Yearbook of the National Society for the Study of Education, pp 182-216.

Brown, P and Elliott, R. (1965). Control of aggression in a nursery school class. J. Experimental Child Psychology, 2, 103-107.

Brown, D., Reschly, D. and Wasserman, H. (1974). Effects of surreptitious modeling upon teacher classroom behaviors. Psychology in the Schools, 11, 366-369.

Brown, J. C., Montgomery, R.and Barclay, J. (1969). An example of psychologist management of teacher reinforcement procedures in the elementary classroom. Psychology in the schools, 6, 336-340.

Burden, R. (1981). Systems theory and its relevance to schools. In: B. Gillham (Ed.): Problem Behaviour in the Secondary School. London: Groom Helm, pp 28-36.

Burns, R. (1982). Self-Concept Development and Education. London: Holt, Rinehart and Winston.

Burt, C. (1965). Factorial studies of personality and their bearing on the work of the teacher. Brit. J. Educ. Psychol., 35, 368-378.

Bushell, D., Wrobel, P.A. and Michaelis, M.L. (1968). Applying group contingencies to the classroom study behavior of preschool children. J. Appl. Beh. Anal. 1, 55-61.

Buys, C.J. (1972). Effects of teacher reinforcement on elementary school pupils' behavior and attitudes. Psychology in the Schools, 9, 278-288.

Campbell, D.T. and Stanley, J.C. (1966). Experimental and Quasi-experimental Designs for Research. Chicago: Rand McNally.

Caplan, G. (1970). The Theory and Practice of Mental Health Consultation. New York: Basic Books.

Carlson, C.S., Arnold, C.R., Becker, W.C. and Madsen, C.H. (1968). The elimination of tantrum behavior of a child in an elementary classroom. Beh. Res. and Ther., 6, 117-119.

Chadwick, B.A. and Day, R.C. (1971). Systematic reinforcement: academic performance of underachieving students. J. Appl. Beh. Analysis, 4, 311-319.

Chazan, M. (1963). Maladjusted pupils: trends in post-war theory and practice. Educ. Res., 6, 29-41.

Chazan, M. (1963). Maladjustment, attainment and sociometric status. Univ. College of Swansea, Faculty of Education Journal, 4-7

Cheatham, R.B. (1968). A study of the effects of group counselling on the self-concept and on reading efficiency of low achieving readers.

Chilman, C.S. (1973). Programmes for disadvantaged parents: some major trends and related research. In: B. M. Caldwell and H.M. Ricciuti (Eds.): Review of Child Development Research. Vol. 3. Chicago: University of Chicago Press.

Christy, P.R. (1975). Does use of tangible rewards with individual children affect peer observers? Journal of Applied Behavior Analysis, 8, 187-196.

Clarizio, H.F. (1976). Towards Positive Classroom Discipline (2nd edition). New York: Wiley.

Clark, C.A. and Walberg, H.J. (1968). The influence of massive rewards on reading achievement in potential urban school dropouts. Amer. Educ. Res. J.,5, 305-310.

Clements, T.H. (1963). A study to compare the effectiveness of individual and group counselling approaches with able under-achievers. Diss. Abs. Int., 24, 1919-1920.

Cobb, J.A. (1972). Relationship of discrete classroom behavior to 4th grade academic achievement. J. Educational Psychology., 63, 74-80.

Cobb, J.A. and Hops, H. (1973). Effects of academic survival skill training on low achieving first grades. J. Educational Research, 67, 108-113.

Cochrane, R.and Sobol, M.P. (1978). Myth and methodology in behaviour therapy research. In: P. Feldman and A Broadhurst (Eds.): Theoretical and Experimental Bases of the Behaviour Therapies. London: John Wiley.

Cohen, J. (1976). A brief comment: evaluating the effectiveness of an unspecified 'casework' treatment in producing change. In: J. Fischer: The Effectiveness of Social Casework. Springfield, Illinois: Charles C. Thomas pp 176-189.

Coleman, J.S. (1966). Equality of Educational Opportunity. Washington: U.S. Government Printing Office.

Colman, A.D. (1975). Operant conditioning and organizational design. In: A.D. Colman and W. H. Bexton: Group Relations Reader. Berkley, California: GREX, pp 327 -337

Combs, T.P. and Slaby, D.A. (1977). Social skills training with children. In: B.B. Lahey and A.E. Kazdin (Eds.): Advances in Clinical Child Psychology (Vol. 1) New York: Plenum pp 161-203

Cone, J.D. (1973). Assessing the effectiveness of programmed generalization. J. Appl. Beh. Anal., 6, 713-718.

Conger, J.C. and Keane, S.P. (1981). Social skills intervention in the treatment of isolated or withdrawn children. Psychol. Bull. 90, 478-495.

Connolly, K. J. (1973). Ethological techniques and the direct observation of behavior. In: P. Mittler (Ed.): Assessment for Learning in the Mentally Handicapped. New York: Churchill-Livingstone.

Cookson, D. (1973). The Junior Eysenck Personality Inventory and teachers' ratings: an investigation of unexpected results. Brit. J. Educ. Psychol., 43, 52-57.

Cooper, M.L., Thomson, C.L., and Baer, D. M. (1970). The experimental modification of teacher attending behavior. J. Appl. Beh. Anal., 3, 153-157.

Copeland, R.E., Brown, R.E. and Hall, R.V. (1974). The effects of principal-implemented techniques on the behaviour of pupils. J. Appl. Beh. Anal,7,77-86.

Copeland, R and Hall, R.V. (1976). Behavior modification in the classroom. In: M. Hersen, R. M. Eisler and P. M. Miller (Eds.): Progress in Behavior Modification (Vo. 3). New York: Academic Press pp45-78.

Cormier, W.H. and Wahler, R.G. (1973). The application of social reinforcement in six junior high school classrooms. In: J. Lang and R. Williams (Eds.): Classroom Management with Adolescents. New York: MSS Information Corporation.

Cossairt, A., Hall, R.U. and Hopkins, B.L. (1973). The effects of experimenter's instructions, feedback and praise on teacher praise and student attending behavior. J. Appl. Beh. Analysis, 6, 89-100.

Cowen, E.L. (1978). Some problems in community program evaluation research. J. Consult. Clin. Psychol., 46, 792-805.

Cowen, E.L., Pederson, A., Babigian, H., 1220, L.D. and Trost, M.A. (1973). Long term follow-up of easily detected vulnerable children. J. Consult. Clin. Psychol., 41, 438-446.

Craig, M and Furst, P.W. (1965) What happens after treatment? A study of potentially delinquent boys. Social Service Review, 39, 165-171

Craighead, L.W., Stunkard, A.J. and O'Brien, R.M. (1981). Behavior therapy and pharmacotherapy for obesity. Arch. Gen. Psychiat., 38, 763-768.

Croft, I.J. and Grygier, I.G. (1956). Social relationships of truants and juvenile delinquents. Human Relations, 9, 439-466.

Cronbach, L. (1975). Beyond the two disciplines of scientific psychology. American Psychologist, 30, 116-127.

Crow, M.L. (1971). A comparison of three group counseling techniques with sixth graders. Elementary School Guidance Counseling, 6, 37-42.

Currie, S.F., Holtzman, W.H. and Swartz, J.D. (1974). Early indicators of personality traits viewed retrospectively. J. Sch. Psychol., 12, 51-59.

Davey, G. (1981). Conditioning principles, behaviourism and behaviour therapy. In: G. Davey (Ed.): Applications of Conditioning Theory. London: Methuen, pp 190-214.

Davidson, H.H. and Greenberg, J.W. (1967). School achievers from a deprived background. (Report, Project No. 2805) Washington, D.C.: Dept. of Health, Educ. & Welfare.

Davie, E.R., Butler, N. and Goldstein, H. (1972). From Birth to Seven. London: Longman.

Davies, B. (1976). Piggies in the middle or 'Who, sir? No, not me, sir'. In: C. Jones-Davies and R.G. Cave (Eds.): The Disruptive Pupil in the Secondary School. London: Ward Lock pp 26-34.

Dearden, R. (1981). Behaviour modification: towards an ethical appraisal. In: K Wheldall (Ed.): The Behaviourist in the Classroom. Birmingham: Educational Review Offset Publications No. 1, pp 113-120.

Deci, E.L. (1971). Effects of externally mediated rewards on intrinsic motivation. J. Pers. & Social Res., 18, 105-115

De Nike, L.D. (1964). The temporal relationship between awareness and performance in verbal conditioning. J. Exp. Psychol., 68, 521-529.

Department of Education and Science, (1978). Special Educational Needs (The Warnock Report). London: H.M.S.O.

Deskin, G. (1968). Effects of differential treatment procedures on reading ability and anxiety level in children with learning difficulties. Diss. Abs. Int. 28B, 3469-3470

Devoge, J.T. and Beck, S. (1978). The therapist-client relationship in behavior therapy. In: M. Hersen, R.M. Eisler and P.M. Miller Progress In Behaviour Modification (Vo. 6). London: Academic Press.

Dickinson, D. (1974). But what happens when you take that reinforcement away? Psychol. in the Schools, 11, 158-160.

Dil, N. and Gotts, E.E. (1971). Improvement of arithmetic self-concept through combined positive reinforcement, peer interaction and sequential curriculum.

Docking, J. W. (1980). Control and Discipline in Schools: Perspectives and Approaches. London: Harper and Row.

Dollinger, S.J. and Thelen, M.H. (1978), Overjustification and children's intrinsic motivation: comparative effects of four rewards. J. Pers. Soc. Psychol., 36, 1259-1269.

Dorfman, E. (1958). Personality outcome of client-centred therapy. Psychol. Monogr. 72 (Whole No. 456).

Douglas, J.W.B. (1967). The Home and the School. London: Panther.

Douglas, V.I., Parry, P,, Marton, P. and Garson, C. (1976). Assessment of a cognitive training program for hyperactive children. J. Abnorm. Child Psychol., 4, 389-410.

Drabman, R.S. and Lahey, B.B. (1974). Feedback in classroom behaviour modification: effects on the target and her classmates. J. Appl. Beh. Anal., 7, 591-598.

Drabman, R.S., Spitalnik, R. and O'Leary, K.D. (1973). Teaching self-control to disruptive children. J. Abnorm. Psychol., 82, 10-16.

Dubey, D.R., Kent, R.N., O'Leary, S.G., Broderick, J.E. and O'Leary, K.D. (1977). Reactions of children and teachers to classroom observers: a series of controlled investigations. Beh. Ther., 8, 887-897.

Duke, D.L. (1979). Environmental influences on classroom management. In: D.L. Duke (Ed.): Classroom Management: The 78th Yearbook of the National Society for the Study of Education. Chicago: Univ. of Chicago Press.

Dulany, D. (1974). On the support of cognitive theory in opposition to behavior theory: a methodological problem. In: W. Weimer and D. Palermo (Eds.): Cognition and the Symbolic Processes. N. J. Hillsdale: Lawrence Erlbaum Associates.

Dunkin, M.J. and Biddle, B.J. (1974). The Study of Teaching. New York: Holt, Rinehart and Winston.

Edwards, A.L. (1975). The Social Desirability Variable in Personality Assessment and Research. New York: Dryden.

Eisenberg, L. (1969). The post-quarter century. Amer. J. Orthopsychiat., 39, 389-401

Erikson, K.T. (1962). Notes on the sociology of deviance. Social Problems, 9, 307-314

Erwin, E. (1978). Behavior Therapy: Scientific, Philosophical and Moral Foundations. Cambridge: Cambridge Univ. Press.

Evers, W.L. and Schwarz, J.C. (1973). Modifying social withdrawal in preschoolers: the effects of filmed modeling and teacher praise. J. Abnorm. Child Psychol., 1 248-256.

Evertson, C. (1980). Differences in instructional activities in high and low achieving junior high classes. Paper presented at the Annual Meeting of the A.E.R.A., Boston.

Eysenck, H.J. (1942). The effects of psychotherapy: an evaluation. J. Consult. Psychol., 16, 319-324.

Eysenck, H.J. (1967). The Biological Basis of Personality. Springfield, Illinois: Charles C. Thomas.

Eysenck, H.J. (1978). An exercise in mega-silliness. American Psychologist, 33, 517.

Eysenck, H.J. and Cookson, D. (1969). Personality in primary school children. 2 - Teachers' ratings. Brit. J. Educ. Psychol., 39, 123-130.

Eysenck, S.B.G. (1965). Manual of the Junior Eysenck Personality Inventory. London: Univ. of London Press.

Eysenck, S.B.J. and Pickup, A.J. (1968). Teachers' ratings of extraversion and neuroticism and children's inventory responses. Brit. J. Educ. Psychol., 38, 94-96.

Feingold, B.D. and Mahoney, M.J. (1975). Reinforcement effects on intrinsic interest: undermining the overjustification hypothesis. Behavior Therapy, 6, 367-377.

Feldhusen, J.F., Thurston, J.R. and Benning, J.J. (1970). Longitudinal analysis of classroom behavior and school achievement. J. Exper. Educ., 38, 4-10.

Ferber, H., Keeley, S.M. and Shemberg, F.M. (1974). Training parents in behavior modification: outcome of and problems encountered in a program after Patterson's work. Beh. Ther., 5, 415-419.

Ferritor, D.E., Buckholdt, D., Hamblin, R.L., and Smith, L. (1972). The noneffects of contingent reinforcement for attending behavior on work accomplished. J. Appl. Beh. Anal., 5, 7-17.

Ferster, C.B. (1967). Arbitrary and natural reinforcement. Psychological Record, 17, 341-347.

Ferster, C.B. and Skinner, B.F. (1957). Schedules of Reinforcement. New York: Appleton-Century-Crofts.

Figg, J. and Ross, A. (1981). Analyzing a school system: a practical exercise. In: B. Gillham (Ed.): Problem Behaviour in the SEcondary School. London: Groom Helm, pp 148-167

Finer Report (1974). Report of the Committee on One-Parent Families. D.H.S.S. Cmnd 5629. London: H.M.S.O.

Fischer, J. (1976). The Effectiveness of Social Casework. Springfield, Illinois : Charles C. Thomas.

Fischer, J. (1978). Does anything work? J. Soc. Service Res. 1, 215-243.

Fisher, B. (1953). Group therapy with retarded readers. J. Educ. Psychol., 6, 354-360.

Fisher, C.W., Filby, N.N. and Marliave, R.S. (1977). Instructional time and student achievement in second grade reading and mathematics. Paper presented at A.E.R.A. Annual Conference, New York.

Fogelman, K. (1978). School attendance, attainment and behaviour. Brit. J. Educ. Psychol., 48, 148-158.

Fontana, D. (1981). Psychology for Teachers. London: Macmillan.

Forehand, R. and Atkeson, B.M. (1977). Generality of treatment effects with parents as therapists: a review of assessment and implementation procedures. Beh. Ther., 8, 575-593.

Forness, S.R. and Esveldt, K.C. (1975). Prediction of high-risk kindergarten children through classroom observation. J. Spec. Educ., 9, 375-387.

Forness, S.R., Guthrie, D and Nihara, K. (1975). Clusters of observable behavior in high-risk kindergarten children. Psychology in the Schools, 12, 263-269.

Foster, S.L. and Ritchey, W.L. (1979). Issues in the assessment of social competence in children. J. Appl. Beh. Anal., 12, 625-658.

Frankel, E. (1946). The social relationship of pre-school children. Sociometry Monograph No. 11, Beacon House, New York, 15-30.

Franks, C.M. and Wilson, G.T. (1978). Annual Review of Behavior Therapy, Theory and Practice. (Vo. 6). New York: Brunner-Mazel.

Franks, C.M. and Wilson, G.T. (1979). Annual Review of Behavior Therapy, Theory and Practice (Vol. 7). New York: Brunner-Mazel.

Franks, C.M. and Wilson, G.T. (1980). Behavioral assessment and methodology. In: C.M. Franks, G.T. Wilson, P.C. Kendall and J.D. Brownell (Eds.): Annual Review of Behavior Therapy (Vol. 8). New York: The Guilford Press.

Frick, T. and Semmel, M.I. (1978). Observer agreement and reliabilities of classroom observational measures. Rev. Educ. Res., 48, 157-184.

Friedman, R.M., Filipczak, J. and Fiordaliso, R. (1977). Within-school generalisation of the Preparation through Responsive Educational (PREP) academic project. Beh. Ther., 8, 986-995.

Freud, S. (1922).Introductory Lectures in Psychoanalysis. London: Allen and Unwin.

Furman, W., Rahe, D and Hartup, W. (1979). Rehabilitation of socially withdrawn preschool children through mixed-age and same-age socialization. Child Dev., 50, 915-922.

Gagne, R.M. (1965). The Conditions of Learning. New York: Holt, Rinehart and Winston.

Galloway, D. (1979). The results of special education. In: D. Galloway and C. Goodwin: Educating Slow-Learning and Maladjusted Children: Integration or Segretation. London: Longman, pp 34-50.

Galloway, D., Ball, T., Blomfield, D. and Seyd, R. (1982). Schools and Disruptive Pupils. London: Longman.

Galloway, D., and Goodwin, C. (1979). Educating Slow-Learning and Maladjusted Children: Integration or Segregation. London: Longman.

Garner, J. and Byng, M. (1973). Inequalities of teacher-pupil contact. Brit. J. Educ. Psychol., 43, 234-243.

Garside, R.F., Birch, H., Scott, D.M., Chambers, S., Kilvin, I., Tweddle, E.G. and Barber, L.M. (1975). Dimensions of temperament in infant school children. J. Child Psychol. Psychiat., 16, 219-231.

Gast, D.L. and Nelson, C.M. (1977). Legal and ethical considerations for the use of time out in special education settings. J. Special Education, 11, 457-467.

Gath, D., Cooper, B., Gattoni, F. and Rockett, D. (1977). Child Guidance and Delinquency in a London Borough. Oxford: Oxford Univ. Press.

Gelfand, D.M. and Hartmann, D.P. (1975). Child Behavior Analysis and Therapy. New York: Pergamon.

Gelso, C.J. (1979). Research in counseling: methodological and professional issues. Counseling Psychologist, 8, 7-36.

Georgiades, N.J. and Phillimore, L. (1975). The myth of the hero-innovator and alternative strategies for organizational change. In: C.C. Kiernan and F.P. Woodford (Eds.): Behaviour Modification with the Severely Retarded. Amsterdam: Associated Scientific Publishers.

Gersten, J.C., Langner, T.S., Eisenberg, J.G., Simcha-Fagan, O.and McCarthy, E.D. (1976). Stability and change in types of behavioral disturbance of children and adolescents. J. Abnorm. Child Psychol., 4, 111-127.

Gibson, H.B. and Hanson, R. (1969). Peer ratings as predictors of school behaviours and delinquency. Brit. J. Soc. and Clin. Psychol., 8, 313-322.

Gillham, B. (1981). Rethinking the problem. In: B. Gillham (Ed.): Problem Behaviour in the Secondary School. London: Groom Helm, pp 9-27.

Glasser, W. (1977). Ten steps to good discipline. Today's Education, 66, 61-63.

Glavin, J.P. (1967). 'Spontaneous' improvement in emotionally disturbed children. Doctoral dissertation: George Peabody College for Teachers.

Glavin, J.P., Quay, H.C. and Werry, J.S. (1971). Behavioral and academic gains of conduct problem children in different classroom settings. Exceptional Children, 37, 441-446.

Glidewell, J. and Swallow, C. (1968). The Prevalence of Maladjustment in Elementary Schools. Chicago: Univ. of Chicago Press.

Glueck, S. and Glueck, E.T. (1964).Potential delinquents can be identified. What next? Brit. J. Criminology, 4, 215-226.

Goldiamond, I. (1974). Toward a constructional approach to social problems. Ethical and constitutional issues raised by applied behavior analysis. Behaviorism, 2, 1-84.

Goldstein, H. (1980). 'Fifteen Thousand Hours': a review of the statistical procedures. J. Child Psychol. Psychiat., 21, 363-369.

Good, T.L. and Brophy, J.E. (1974). Changing teacher and student behavior: an empirical investigation. J. Educ. Psychol., 66, 390-405.

Good, T.L. and Brophy, J.E. (1978). Looking in Classrooms (2nd edition). New York: Harper and Row.

Good, T.L., Schmidt, L., Peck, R and Williams, D. (1969). Listening to Teachers. Report Series No. 34. Rsearch and Development Centre for Teacher Education. Univ. of Texas.

Gordon, T. (1974). Teacher Effectiveness Training. New York: Peter H. Wyden.

Gottman, J. (1977). The effects of a modeling film on social isolation in preschool children: a methodological investigation. J. Abnorm. Child Psychol., 5, 69-78.

Gottman, J., Gonso, J. and Schuler, P. (1976). Teaching social skills to isolated children. J. Abnorm. Child Psychol., 4, 179-197.

Gottman, J.M. and Markman, H.J. (1978). Experimental designs in psychotherapy research. In: S.L. Garfield and A.E. Bergin (Eds.): Handbook of Psychotherapy and Behavior Change: An Empirical Analysis. New York: Wiley, pp 23-62.

Graham, P.J. (1979). Epidemiological studies. In: H. C. Quay and J. S. Werry (EDs.): Psychopathological Disorders of Childhood (2nd edition). New York: Wiley, pp 185-209.

Graham, P. and Rutter, M. (1973). Psychiatric disorder in the young adolescent: a follow-up study. Procced. Roy. Soc. Med., 66, 1226-1229.

Graubard, P.S., Rosenberg, H and Miller, M.B. (1971). Student application of behavior modification to teachers and environments or ecological approaches to social deviancy. In: E.A. Ramp and B.L. Hopkins (Eds.): A New Direction for Education: Behavior Analysis. Lawrence, Kansas: Support and Development Center for Follow Through.

Graziano, A.M. (1977). Parents as behavior therapists. In: M. Hersen, R.M. Eisler and P.M. Miller (Eds.): Progress in Behavior Modification Vol. 4. New York: Academic Press.

Greenspoon, J. (1962). Verbal conditioning and clinical psychology. In: A.J. Bachrach (Ed.): Experimental Foundations of Clinical Psychology. New York: Basic Books.

Greenwood, C.R., Hops, H., Delquadri, J. and Guild, J. (1974). Group contingencies for group consequences in classroom Management: a further analysis. J. Appl. Beh. Anal., 7, 413-425.

Greenwood, C.R., Hops, H and Walker, H.M. (1977). The program for academic survival skills (PASS): effects on student behavior and achievement. J. School Psychology, 15, 25.35.

Gronlund, N.E. (1953). Generality of sociometric status over criteria in the measurement of social acceptability; Elementary School Journal , LVI, 173-176.

Gronlund, N.E.(1970). Sociometry in the Classroom. Bath: Cedric Chivers Ltd.

Grossman, B. and Wrighter, J. (1948). The relationship between selection-rejection and intelligence, social status, and personality among sixth-grade children. Sociometry, 11, 346-355.

Guilford, J.P . Christensen, P.R., Taaffe, G. and Wilson, R.C. (1962). Ratings should be scrutinized. Educ. Psychol. Measmt .., 22, 439-447.

Gurin, G., Veroff, J. and Feld, S. (1960). Americans View Their Mental Health: A Nationwide Survey. New York: Basic Books.

Gurney, P.W. (1979). The use of behaviour modification to increase the level of self-esteem in maladjusted boys. Unpublished doctoral thesis: Univ. of Exeter.

Hall, R.V. and Copeland, R.D. (1972). The Responsive Teaching Model: A first step in shaping school personnel as behaviour modification specialists. In: F.W. Clask, D.R. Evans and L.A. Hamerlynck (Eds.): Implementing Behavioural Programs for Schools and Clinics. Champaign, Illinois; Research Press.

Hall, R.V., Copeland, R and Clark, M. (1976). Management strategies for teachers and parents: Responsive teaching. In: N.G. Haring and R.L. Schiefelbusch (Eds.): Teaching Special Children. New York: McGraw Hill.

Hall, R.V., Fox, R., Willard, D., Goldsmith, L., Emerson, M., Owen, M., Davis, F., and Porcia, E. (1971). The teacher as observer and experimenter in the modification of disputing and talking out behaviors. J. Appl. Beh. Analysis, 4, 141-149.

Hall, R.V., Lund, D. and Jackson, D. (1968). Effects of teacher attention on study behavior. J. of Applied Beh. Analysis,1,1-12.

Hall, R.V., Panyan, M., Rabon, D. and Broden, M. (1968). Instructing beginning teachers in reinforcement procedures which improve classroom control. J. Appl. Beh. Anal., 1, 315-322.

Hallam, R.S. (1976). The Eysenck personality scales: stability and change after therapy. Beh. Res. and Ther., 14, 369-372.

Hallett, S.E. (1981). Some thoughts on subjective private events. Behaviour Analysis, 3, 26-36.

Hallworth, H.J. (1966). Perception of children's personalities by experienced teachers. Educ. Rev., 19, 3-12.

Hampe, I.E., Noble, H., Miller, L.C., and Barrett, C.L. (1973). Phobic children one and two years post-treatment. J. Abnorm. Psychol., 82, 446-453.

Hanko, G. (1981). Staff support and training groups for teachers. New Growth, 1, 23-28.

Hanley, E.M. and Perelman, P.F. (1971). Research resulting from a model cities program designed to train paraprofessionals to aid teachers in elementary school classrooms. In: E.A. Ramp and B.L. Hopkins (Eds.): A New Direction for Education: Behavior Analysis. Lawrence, Kansas: Support and Development Center for Follow Through.

Hargreaves, D.H. (1967). Social Relationships in a Secondary School. London: Routledge and Kegan Paul.

Hargreaves, D.H. (1978). Deviance: the interactionist approach. In: B. Gillham (Ed.): Reconstructing Educational Psychology. London: Groom-Helm, pp 67-81.

Hargreaves, D.H. (1978). What teaching does to teachers. New Society, 43, 540-542.

Hargreaves, D.H. (1980). Review of 'Fifteen Thousand Hours'. Brit. J. Sociol. Educ., 1, 211-216.

Hargreaves, D.H., Hester, S.K. and Mellor, F.J. (1975). Deviance in Classrooms. London: Routledge and Kegan Paul.

Haring, N.G. (1968). Behavior modification and special education. The experimental unit - University of Washington. U.C.L.A.: 4th Annual Engineering Conference, School of Education.

Haring, N.G., Hayden, A.H. and Nolen, P.A. (1969). Accelerating appropriate behaviors of children in a Head Start program. Exceptional Children, 35, 773-786.

Harnischfeger, A. and Wiley, D.E. (1975). Teaching/learning processes in elementary school: a synoptic view. Studies of Education Processes, No.9, University of Chicago.

Harper, D. (1965). Some correlates of sociometric status among adolescent boys: Report to the Study Group on Research Methods in Child Development. Unpublished: London School of Economics.

Harper, D.G. (1968). The reliability of measures of sociometric acceptance and rejection. Sociometry, 31, 219-227.

Harre, R. and Secord, P.F. (1972). The Explanation of Social Behaviour. London: Blackwell.

Harris, F.R., Johnston, M.K., Kellem, C.S. and Wolf, M.M. (1964). Effects of social reinforcement on regressed crawling of a nursery school child. J. Educational Psychology, 55, 35-41.

Harris, V.W. and Sherman, J.A. (1974). Homework assignments, consequences and classroom performance in social studies and mathematics. J. Appl. Beh. Analysis, 7, 505-519.

Harrop, L.A. (1974). A behavioural workshop for the management of classroom problems. British Journal of In-service Education, 1, 1, 47-50.

Harrop, L.A. (1983). Behaviour Modification in the Classroom. London: Hodder and Stoughton.

Hart, B.M., Allen, K.E., Buell, J.S., Harris, F.R. and Wolf, M.M. (1964). Effects of social reinforcement on operant crying. J. Experimental Child Psychology, 1, 145-153.

Hart, B.M., Reynolds, N.J., Baer, D.M., Branley, E.R. and Harris, F.R. (1968). Effect of contingent and non-contingent social reinforcement on the co-operative play of a preschool child. J. Appl. Beh. Analysis, 1, 73-76.

Hartman,, D.P. (1977). Considerations in the choice of interobserver reliability estimates. J. Appl. Beh. Anal., 10, 103-116.

Hartmann, D.P. and Atkinson, C. (1973). Having your cake and eating it too: a note on some apparent contradictions between therapeutic achievements and design requirements in N=1 studies. Beh. Ther., 4, 589-591.

Hartmann, D.P., Roper, B.L. and Gelfand, D.M. (1977). An evaluation of alternative modes of child psychotherapy. In: B.B. Lahey and A.E. Kazdin (Eds.): Advances in Clinical Child Psychology, 1. New York: Plenum, pp 1-46.

Hartmann, D.P. and Wood, D.D. (1982). Observational methods. In: A.S. Bellack, M. Hersen and A.E. Kazdin (Eds.): International Handbook of Behavior Modification and Therapy. New York: Plenum Press, pp 109-138.

Hartup, W.W. (1970). Peer interactions and social organisation. In: P.H. Mussen (Ed.): Carmichael's Manual of Child Psychology, Vol. 2, 361-456. London: John Wiley.

Hartup, W.W., Glazer, J.A. and Charlesworth, R. (1967). Peer reinforcement and sociometric status. Child Dev., 38, 1017-1024.

Hasazi, J.E. and Hasazi, S.E. (1972). Effects of teacher attention on digit reversal behavior in an elementary school child. J. Appl. Beh. Anal, 5, 157-162.

Hawkins, P.J. (1974). Training teachers in classroom management. Behaviour Modification Newsletter, Issue 5, 7-14.

Hay, W.M., Hay, L, and Nelson, R.O. (1977). Direct and collateral changes in on-task and academic behavior resulting from on-task versus academic contingencies. Behavior Therapy, 8, 431-441.

Heaton, R.C., Safer, D.J., Allen, R.P., Spinnato, N.C. and Prumo, F.M. (1976). A motivational environment for behaviourally deviant junior high school students. J. Abnormal Child Psychology, 4, 263-275.

Heinicke, C.M. (1969). Frequency of psychotherapeutic sessions as a factor affecting outcome: analysis of clinical ratings and test results. J. Abnorm. Pscyhol,, 74, 533-560.

Herbert, G.W. (1974). Teachers' ratings of classroom behaviour: factorial structure. Brit. J. Educ. Psychol., 44, 233-240.

Herbert, M. (1978). Conduct Disorders of Childhood and Adolescence. London: Wiley.

Hersen, M. (1981). Complex problems require complex solutions. Behavior Therapy, 12, 15-29.

Hersen, M. and Barlow, A.S. (Eds.) (1976). Single-Case Experimental Designs: Strategies for Studying Behavior Change. New York: Pergamon.

Hewett, F.M., Taylor, F.D., Artuso, A.A. (1969). The Santa Monica Project: Evaluation of an engineered classroom design with emotionally disturbed children. Exceptional Children, 35, 523-529.

Hilgard, E.R. and Bower, G. (1966). Theories of Learning (3rd ed.) New York: Appleton-Century-Crofts.

H.M.I. (1979). Aspects of secondary education in England. London: H.M.S.O.

Hogg, J. (1976). The experimental analysis of retarded behaviour and its relation to normal development. In: P. Feldman and A. Broadhurst (Eds.): Theoretical and Experimental Bases of the Behaviour Therapies. London: Wiley, pp 151-189.

Hollins, T.H.B. (1955). Teachers' attitudes to children's behaviour. M. Ed. thesis: Univ. of Manchester.

Honig, W.K. and Stadden, J.E.R. (Eds.) (1977). Handbook of Operant Behavior. Englewood Cliffs, New Jersey: Prentice-Hall.

Hopkins, B.L. and Hermann, J.A. (1977). Evaluating interobserver reliability of interval data. J. Appl. Beh. Anal., 10, 121-126.

Hops, H. and Cobb, J.A. (1973). Survival behaviors in the educational setting: their implications for research and intervention. In: L.A. Hamerlynck, L.C. Handy and E. J.Mash (Eds.): Behavior Change: Methodology, Concepts and Practice. Champaign, Illinois: Research Press, pp 193-208.

Hopson, B. (1982). Counselling and helping. In: J. Hall (Ed.): Psychology for Nurses and Health Visitors. London: Macmillan, pp 250-266.

Horowitz, F.D. (1962). The relationship of anxiety, self-concept and sociometric status among 4th, 5th and 6th grade children. J. Abnorm. and Soc. Psychol., 65, 212-214.

Horton, G.O. (1975). Generalization of teacher behaviour as a function of subject matter specific discrimination training. J. Appl. Beh. Anal., 8, 311.320.

House, R.M. (1971). The effects of nondirective group play therapy on the sociometric status and self-concept of selected second grade children. Diss. Abs. Int., 31A, 2684.

Hubbert, A.K. (1970). Effect of group counselling and behavior modification on attentive behavior of first grade students. Diss. Abs. Int., 30A, 3727.

Hudson, W.W. (1976). Special problems in the assessment of growth and deterioration. In: J. Fisher: The Effectiveness of Social Casework. Springfield, Illinois: Charles C. Thomas, pp 197-224.

Hugo, M.J. (1970). The effects of group counseling on self-concept and behavior of elementary school children. Diss. Abs. Int., 30A, 3728.

Hume, K.E. (1967). A counseling service project for grades one through four. Diss. Abs. Int., 27A, 4130.

Humphrey, J. (1966). Behavioural therapy with children: an experimental evaluation. Unpublished doctoral dissertation: Univ. of London.

Ingham, J.G. (1966). Change in MPI scores in neurotic patients: a three year follow-up. Brit. J. Psychiat., 112, 931-939.

Iwata, B.A. and Bailey, J.S. (1974). Reward versus cost token systems: an analysis of the effects on students and teacher. J. Appl. Beh. Anal., 7, 567-576.

Izard, C.E. (1959). Personality correlates of sociometric status. J. Appl. Psychol., 43, 89-93.

Jackson, P.W. (1968). Life in Classrooms. New York: Holt, Rinehart & Winston.

Jackson, P.W. and Lahaderne, H.M. (1967). Inequalities of teacher-pupil contact. Psychology in the Schools, 4, 201-211.

Jason, L.A. and Ferone, L. (1978). Behavioral versus process consultation interventions in school settings. Amer. J. Community Psychol., 6,531-544.

Jason, L.A.,Ferone, L. and Anderegg, T. (1979). Evaluating ecological, behavioral and process consultation interventions. J. School Psychol., 17, 103-115.

Jencks, C. (1972). Inequality: A Reassessment of the Effect of Family and Schooling in America. New York: Basic Books.

Johnson, C.A. and Katz, R.C. (1973). Using parents as change agents for their children: a review. J. Child Psychology and Psychiatry, 14, 181-200.

Johnson, S.M. and Bolstad, O.D. (1973). Methodological issues in naturalistic observation: some problems and solutions for field research. In: L.A. Hamerlynck, L.C. Handy and G.J. Mash (Eds.): Behavior Change: Methodology, Concepts and Practice. Champaign, Illinois: Research Press.

Johnson, S.M. and Christensen, A. (1975). Multiple criteria follow-up of behavior modification with families. J. Abnorm. Child Psychol., 3, 135-154.

Johnson, S.M. and Eyberg, S. (1975). Evaluating outcome data: a reply to Gordon. J. Consult. Clin. Psychol., 43, 917-919.

Johnston, M.S., Kelley, C.S., Harris, F.R. and Wolf, M.M. (1966). An application of reinforcement principles to development of motor skills of a young child. Child Development, 37, 379-387.

Johnston, M.S., Kelley, C.S., Harris, F.R., Wolf, M.M. and Baer, D.M. (1964). Effects of positive social reinforcement on isolate behavior of a nursery school child. Unpublished manuscript: Univ. of Washington.

Jones, R.R., Reid, J.B. and Patterson, G.R. (1975). Naturalistic observation in clinical assessment. In: P. McReynolds (Ed.): Advances in Psychological Assessment, Vol. 3. San Francisco: Jossey-Bass.

Jones, R.R., Vaught, R.S. and Weinrott, M. (1977). Time series analysis in operant research. J. Appl. Beh. Anal., 10, 151-166.

Jones, R.T. and Kazdin, A.E. (1975). Programming response maintenance after withdrawing token reinforcement. Beh. Ther., 6, 153-164.

Karoly, D. (1980). Operant methods. In: F.H. Kanfer and A. P. Goldstein (Eds.): Helping People Change (2nd ed.) London: Pergamon, pp 210-247.

Katz, D. and Kahn, R.L. (1966). The Social Psychology of Organizations. New York: Wiley.

Kauffman, J.M., Nussen, J.L. and McGee, C.S. (1977). Follow-up in classroom behavior modification: survey and discussion. J. School Psychology, 15, 343-348.

Kaufman, K.F. and O'Leary, K.D. (1972). Reward, cost and self-evaluation procedures for disruptive adolescents in a psychiatric hospital school. J. Appl. Beh. Analysis, 5, 293-310.

Kazdin, A.E. (1973). Methodological and assessment considerations in evaluating reinforcement programs in applied settings. J. Appl. BEh. Anal., 6, 517-531.

Kazdin, A.E. (1977a). Vicarious reinforcement and direction of behavior change in the classroom. Behavior Therapy, 8, 57-63.

Kazdin, A.E. (1977b). The Token Economy. New York: Plenum.

Kazdin, A.E. (1979). Artifact, bias, and complexity of assessment: the ABCs of reliability. J. Appl. Beh. Anal., 10, 141-150.

Kazdin, A.E. (1981a). Behavior modification in education: contributions and limitations. Developmental Review, 1, 34-57.

Kazdin, A.E. (1981b). Drawing valid inferences from case studies. J. Consult. Clin. Psychol., 49, 183-192.

Kazdin, A.E. (1981c). Behavioral observation. In: M. Hersen and A.S. Bellack (Eds.) Behavioral Assessment: A Practical Handbook. New York: Pergamon, pp 101-124.

Kazdin, A.E. (1982). Symptom substitution, generalization and response covariation: implications for psychotherapy outcome. Psychol. Bull., 91,. 349-365.

Kazdin, A.E. and Bootzin, R.R. (1972). The token economy: an evaluative review. J. Appl. Beh. Anal., 5, 343-372.

Kazdin, A.E. and Kopel, S.A. (1975). On resolving ambiguities of the multiple baseline design: problems and recommendations. Beh. Ther., 6, 601-608.

Kazdin, A.E. and Moyer, W. (1976). Training teachers to use behavior modification. In: S. Yen and R. McIntire (Eds.): Teaching Behavior Modification. Kalamazoo, Michigan: Behaviordelia.

Kazdin, A.E. and Wilcoxon, L.A. (1976). Systematic desensitization and nonspecific treatment effects: a methodological evaluation. Psychol. Bull., 83, 729-758.

Kazdin, A.E. and Wilson, G.T. (1978). Evaluation of Behavior Therapy. Cambridge, Mass,: Ballinger.

Keeley.S.M., Shemberg, K.M. and Carbonnel, J. (1976). Operant clinical interventions: behavior management or beyond? Where are the data? Behavior Therapy, 7, 292-305.

Keller, M. F. and Carlson, P.M. (1974). The use of symbolic modelling to promote social skills in preschool children with low levels of social responsiveness. Child Dev., 45, 912-919.

Kelly, T.J., Bullock, L.M. and Dykes, M.K. (1977). Behavioral disorders: teachers' perceptions. Exceptional Children, 43, 316-318.

Kendall, P.C. and Hollon, S.D. (Eds.) (1979). Cognitive and Behavioral Interventions: Theory, Research and Procedures. New York: Academic Press.

Kent, R.N. (1976). A methodological critique of 'Intervention for boys with conduct problems'. J. Consult. Clin. Psychol., 44, 297-299.

Kent, R.N. and Foster, S.L. (1977). Direct observational procedures: methodological issues in naturalistic settings. In: A.R. Ciminero, K.S. Calhoun and H.E. Adams, (Eds.): Handbook of Behavioral Assessment. New York: Wiley.

Kent, R.N., Kanowitz, J., O'Leary, K.D. and Cheiken, M. (1977). Observer reliability as a function of circumstances of assessment. J. Appl. Beh. Anal., 10, 317-324.

Kent, R.N., and O'Leary, K.D. (1976). A controlled evaluation of behavior modification with conduct problem children. J. Consulting and Clinical Psychology, 44, 586-596.

Kent, R.N., O'Leary, K.D. and Broderick, J. (1976). A manual of intervention strategies in the school: rationale, assessment and procedures. New York: State University of New York at Stony Brook.

Kent, R.N., O'Leary, K.D., Diament, C., and Dietz, A. (1974). Expectation biases in observational evaluation of therapeutic change. J. Consult. Clin. Psychol., 42 ,774-780.

Khlief, B.B. (1976). Role distance of classroom teachers of slow learners. J. Res. Dev. Educ., 9, 69-73.

Kiernan, C.C. (1974). Methodology of behavior modification. In: C.C. Kiernan and F.P. Woodford (Eds.): Behavior Modification with the Severely Retarded. Amsterdam: Associated Scientific Publishers.

Kiesler, D.J. (1966). Some myths of psychotherapy research and the search for a paradigm. Psychol. Bull., 65, 110-136.

Kiesler, D.J. (1971). Experimental designs in psychotherapy research. In: A.E. Bergin and S.L. Garfield (Eds.): Handbook of Psychotherapy and Behavior Change. New York: Wiley.

Kiesler, D.J. (1981). Empirical clinical psychology: myth or reality? J. Consult. Clin. Psychol., 149, 212-215.

Kimble, G.A. (1961). Hilgard and Marquis' Conditioning and Learning. New York: Appleton Century Crofts.

Kirby, F.D. and Shields, F. (1972). Modification of arithmetic response rate and attending behavior in a seventh-grade student. J. Appl. Beh. Anal., 5, 79-84.

Klein, A.F. (1959). Social work in non-social work settings. Social Work, 4, 92-97.

Knight, M.F., Hasazi, S.E. and McNeil, M.E. (1971). A home-based program for the development of reading skills for preschoolers. In: E.A. Ramp and B.L. Hopkins (Eds.): A New Direction for Education: Behavior Analysis. Lawrence, Kansas: Support and Development Center for Follow Through.

Kolvin, I., Garside, R.F., Hulbert, C.M., Van Der Spuy, H., Wolstenholme, F and Wrate, R. (1973). Evaluation of psychiatric services for children in England and Wales. In: J.K. Wing and H. Hafner (Eds.): Roots of Evaluation. London: Oxford Univ. Press, pp 131-162.

Kolvin, I., Garside, R.F., Nicol, A.R., Macmillan, A., Wolstenholme, F. and Leitch, I.M.(1981). Help Starts Here: The Maladjusted Child in the Ordinary School. London: Tavistock.

Kolvin, I., Garside, r.F., Nicol, A.R., Leitch, I.M., and Macmillan, A. (1977). Screening school children for high risk of emotional and educational disorder. Brit. J. Psychiatry, 131, 192-206.

Kolvin, I., Wolff, S., Barber, L.M., Tweddle, E.G., Garside, R.F., Scott, D.M. and Chambers, S. (1975). Dimensions of behaviour in infant school children. Brit. J. Psychiat., 126, 114-126.

Kounin, J. (1970). Discipline and group management in classrooms. New York: Holt, Rinehart and Winston.

Kranzler, G.D., Mayer, G.R., Dyer, C.O., and Munger, P.F. (1966). Counseling with elementary school children: an experimental study. Personnel and Guidance Journal, 5, 944-949.

Kubany, E.S., Sloggett, B.B. and Ogata, R.F. (1974). Training teachers in classroom management skills. Amer. J. Comm. Psychol., 2, 139-149.

Kubany, E.S., Weiss, L.E. and Sloggett, B.B. (1971). The good behavior clock: a reinforcement/time-out procedure for reducing disruptive classroom behavior. J. Behavior Therapy and Experimental Psychiatry, 2, 173-179.

Kuhn, T. (1962). The Structure of Scientific Revolutions. Chicago: The Univ. of Chicago Press.

Kuypers, D.S., Becker, W.C., and O'Leary, K.D. (1968). How to make a token system fail. Exceptional Children, 11, 101-108.

Kyriacou, C. and Sutcliffe, J. (1978). A model of teacher stress. Educational Studies, 4, 1-6.

Lahaderne, H.M. (1968). Attitudinal and intellectual correlates of attention: a study of fourth-grade classrooms. J. Educ. Psychology, 59, 320-324

Lahey, B.B. (1977). Research on the role of reinforcement in reading instruction: some measurement and methodological deficiencies. Corr. & Soc. Psychiat. & J. Beh. Tech. Meth. & Ther.,23, 27-32.

Lahey, B.B., Hobbs, S.A., Kupfer, D.L. and Delamater, A. (1978). Current perspectives on hyperactivity and learning disabilities. IN: B.B. Lahey (Ed.): Behavior Therapy with Hyperactive and Learning Disabled Children. Oxford: Oxford Univ. Press.

Lahey, B.B., McNees, M.P. and Brown, C.C. (1973) Modification of deficits in reading for comprehension. J. Appl. Beh. Anal., 6, 475-480.

Lapouse, R and Monk, M.A. (1959). Fears and worries in a representative sample of children. Amer. J. Orthopsychiat., 29, 803-818.

Laslett, R. (1977). Educating Maladjusted Children. London: Crosby Lockwood Staples.

Lates, B.J., Egner, A.N. and McKenzie, H.S. (1971). Behavior analysis of the academic and social behavior of first grade children. In: E. A. Ramp & B L. Hopkins (Eds.): A New Direction for Education. Lawrence, Kansas: Support and Development Center for Follow-Trhough.

Lavik, N. (1977). Urban-rural differences in rates of disorder: a comparative psychiatric population study of Norwegian adolescents. In: P.J. Graham (Ed.): Epidemiological Approaches in Child Psychiatry. New York: Academic Press.

Leach, D. (1977). Teachers' perceptions and 'problem' pupils. Educational Review, 29, 188-203.

Lehrman, L.J., Sirluck, H., Black, B.J. and Glick, S.J. (1949). Success and failure of treatment of children in the Child Guidance Clinics of the Jewish Board of Guardians. Research Monograph No. 1, New York: Jewish Board of Guardians.

Leitenberg, H. (1973). The use of single-case methodology in psychotherapy research. J. Abnorm. Psychol., 82, 87-101.

Leitenberg, H. and Callahan, E.J. (1973). Reinforced practice and reduction of different kinds of fears in adults and children. Beh. Res. & Ther., 11, 19-30.

Lepper, M.R., Greene, D., & Nisbett, R.E. (1973). Undermining children's intrinsic interest with extrinsic rewards: a test of the 'overjustification' hypothesis. J. Pers. & Social Psychol., 28, 129-137

Leslie, S.A. (1974). Psychiatric disorder in the young adolescents of an industrial town. Brit. J. Psychiat., 125, 113-124.

Levine, F.M. and Fasnacht, G. (1974). Token rewards may lead to token learning. American Psychologist, 29, 816-820.

Levitt, E.E. (1957). The results of psychotherapy with children an evaluation. J. Consult. Psychol., 21, 189-186

Levitt, E.E. (1958). A comparative judgemental study of 'defection' from treatment at a child guidance clinic. J. Clin. Psychol., 14, 429-432.

Levitt, E.E. (1963). Psychotherapy with children: a further evaluation. Beh. Res. & Ther., 60, 326-329.

Levitt, E.E. (1971). Research on psychotherapy with children. In: A.E. Bergin and S.L. Garfield (Eds.): Handbook of Psychotherapy and Behavior Change. New York: John Wiley.

Levitt, E.E., Beiser, H.R. and Robertson, R.E. (1959). A follow-up evaluation of cases treated at a community child guidance clinic. Amer. J. Psychiat., 29, 337-347.

Liberman, R.P. (1980). Review of 'Psychosocial Treatment for chronic mental patients' by G.L. Paul and R. J. Lentz. J. Appl. Beh, Ana., 13, 367-372.

Liberman, R.P., Ferris, C., Salgado, P., and Salgado, J. (1975). Replication of the Achievment Place model in California. J. Appl. Beh. Anal., 8, 287-299.

Lieberman, M.A., Yalom, I.D. and Miles, M.B. (1973) Encounter Groups: First Facts. New York: Basic Books.

Lippitt, R. and Gold, M. (1959). Classroom social structure as a mental health problem. J. Social Issues, 15, 40-49.

Litow, L. and Pumroy, D.K. (1975). A brief review of classroom group-oriented contingencies. Journal of Applied Behavior Analysis, 8, 341-347.

Long, J.L., Morse, C.M., and Newman, R.G. (1971). Conflict in the Classroom. Belmont, California: Wadsworth.

Lord, F.M., (1960). Large-sample covariance analysis when the control variable is fallible. J. Amer. Stat. Assoc., 55, 307-321.

Lovaas, O.I. (1969). Behavior Modification: Teaching Language to Psychotic Children. New York: Appleton-Century-Crofts.

Love, L.R., Kaswan, J. and Bugental, D.E. (1972) Differential effectiveness of of three clinical interventions for different socio-economic groupings. J. Consult. Clin, Psychol., 39, 347-360.

Lovell, K. (1958). Educational Psychology and Children. London: Univ. of London Press.

Lowenstein, L.F. (1975). Violent and Disruptive Behaviour in Schools. Hemel Hempstead: National Association of Schoolmasters.

Lyon, S. (1977). Teacher non-verbal behavior related to perceived pupil social-personal attributes. J. Learning Disabilities, 10, 173-177.

McAllister, L.W., Stachowiak, J.G., Baer, D.M. and Conderman, L. (1969). The application of operant conditioning techniques in a secondary school classroom. J. Appl. Beh. Analysis, 2, 277-285.

McBrien, R.J. and Nelson, R.I. (1972). Experimental group strategies with primary grade children. Elementary School Guidance and Counselling, 6, 170-174.

McCaffrey, I. and Cumming, J. (1969). Persistence of emotional disturbances reported among 2nd and 4th grade children. In: H. Dupont (Ed.): Educating Emotionally Disturbed Children. New York: Holt, Rinehart & Winston.

McCandless, B.R., Castaneda, A., and Palermo, D.S. (1956). Anxiety and children and social status. Child Dev., 27, 385-391.

McCollum, P.S. and Anderson, R.P. (1974). Group counseling with reading disabled children. J. Counsel. Psychol., 21, 150-155.

McDonald, F.H. (1975). Research on teaching and its implications for policy making: Report of Phase II of beginning teacher evaluation study. Paper presented at Conference on Research on Teacher Effects. University of Texas.

McDonald, S. (1973). The kibitz dimension in teacher consultation. In: R.D. Klein, W. G. Hepkiewicz and A.H. Roden: Behavior Modification in Educational Settings. Springfield, Illinois: Thomas.

Macfarlane, J.W., Allen, L. and Honzik, M.P. (1954). A Developmental Study of the Behavior Problems of Normal Children Between Twenty One Months and Fourteen Years. Berkeley and Los Angeles: Univ. of California Press.

McKenzie, H.S., Egner, A.N., Knight, M.F., Perelman, P.F., Schneider, B.M. and Garvin, J.S. (1970). Training consulting teachers to assist elementary teachers in the management of handicapped children. Exceptional Children, 37, 137-143.

McKeown, D., Adams, H.E. and Forehand, R. (1975). Generalization to the classroom of principles of behavior modification taught to teachers. Beh. Res. & Ther., 13, 85-92.

McKinney, J.D., Mason, J., Perkerson, K. and Clifford, M. (1975). Relationship between classroom behavior and academic achievement. J. Educ. Psychol., 67, 198-203.

McLaughlin, T.F. (1975). The applicability of token reinforcement systems in public school systems. Psychology in the Schools, 12, 84-89

Macmillan, A. and Kolvin, I. (1977). Behaviour modification in educational settings: a guide for teachers. Journal of the Association of Workers for Maladjusted Children, 5, 2-18.

Macmillan, A., Kolvin, I., Garside, R.F., Nicol, A.R. and Leitch, I.M. (1980). A multiple criterion screen for identifying secondary schoolchildren with psychiatric disorder. Psychological Medicine, 10, 265-276.

Macmillan, A., Walker, L., Garside, R.F., Kolvin, I., Leitch, I. and Nicol, A.R. (1978). The development and application of sociometric techniques for the identification of isolated and rejected children. Journal of the Association of Workers for Maladjusted Children, 6, 58-74.

Macmillan, D.L. (1973). Behavior Modification in Education. New York: Macmillan.

Macmillan, D.L., Forness, S.R. and Trumbull, B.M. (1973). The role of punishment in the classroom. Exceptional Children, 40, 85-97.

Macmillan, D.L. and Morrison, G.M. (1980). Evolution of behaviorism from the laboratory to special education settings. In: B.K. Keogh (Ed.): Advances in Special Education, Vol. 2, Perspectives on Application. New York: JAI Press, pp1-28.

McNamara, E. (1979). The use of self-recording in behaviour modification in a secondary school. Behavioural Psychotherapy, 7, 57-66.

McNamara, E. and Harrop, L.A. (1979). Behaviour modification in the secondary school: a cautionary tale. B.P.S.: Division of Educational & Child Psychology Occasional Papers 3, 38-41.

McNamara, J.R. (1971). Teachers and students as a source for behavior modification in the classroom. Behavior Therapy, 2, 205-213

McNamara, J.R. (1975). Ways by which outcome measures influence outcomes in classroom behaviour modification research. J. School. Psychol., 13, 104-113

Madsen, C.H., Becker, W.C., and Thomas, D.R. (1968). Rules, praise and ignoring: elements of elementary classroom control. J. Appl. Beh. Anal., 1, 139-150.

Madsen, C.H. and Madsen, C.K. (1973). Teaching/Discipline: Behavioral Principles toward a Positive Approach. Boston: Allyn & Bacon.

Madsen, C.H., Madsen, C.K., Saudargas, R.A., Hammond, W.R., Smith, J.B. and Edgar, D.E. (1970). Classroom Raid (Rules, approval, ignore, disapproval); a co-operative approach for professionals and volunteers. J. School Psychol., 8, (3), 180-185.

Mahoney, M.J. (1979). Cognitive skills and athletic performance. In: P.C. Kendall and S.D. Hollon (Eds.): Cognitive-Behavioural Interventions: Theory, Research and Procedures. New York: Academic Press, pp 423-444.

Main, G.C. and Munro, B.C. (1975). A token reinforcement program in a public junior high school. J. Appl. Beh. Anal., 10, 93.

Mann, P.H. (1968). The effect of group counseling on educable mentally retarded boys. Diss. Abs. Int. 28A, 3467.

Mannino, F.V. and Shore, M.F. (1975). The effects of consultation: a review of empirical studies. Amer. J. Commun. Psychol., 3, 1-21.

Marholin, D.II, Siegel, L.J. and Phillips, D. (1976). Treatment and transfer: a search for empirical procedures. In: M. Hersen, R.M. Eisler and P. M. Miller (Eds.): Progress in Behavior Modification (Vol. 3). New York: Academic Press.

Marland, M. (1975). The Craft of the Classroom. London: Heineman.

Marlowe, R.H., Madsen, C.H., Bowen, C.E., Reardon,R. and Logue, P.E. (1978). Severe classroom behavior problems: teachers or counsellors. J. of Appl. Beh. Anal., 11, 53-66.

Marsh, P., Rosser, E. and Harre, R. (1978). The Rules of Disorder. London: Routledge and Kegan Paul.

Marshall, T.F. and Rose, G. An experimental evaluation of school social work. Brit. J. Guid. & Counselling, 3, 2-13.

Mash, E.J. (1975). Behavior modification and methodology: a developmental perspective. Unpublished manuscript: Univ. of Calgary.

Mash, E.J. and Terdal, L.G. (1977). After the dance is over: some issues and suggestions for follow-up assessment in behavior therapy. Psychol. Bull., 41, 1287-1308.

Medway, F.J. (1979). How effective is school consultation: a review of recent research. J. Sch. Psychol., 17, 275-282.

Meehl, P.E. (1978). Theoretical risks and tabular risks: Sir Karl, Sir Ronald and the slow progress of soft psychology. J. Consult. Clin. Psychol., 46, 806-834.

Meichenbaum, D. (1977). Cognitive-Behaviour Modification: An Integrative Approach. New York: Plenum Press.

Meichenbaum, D.H., Bowers, K. and Ross, R.R. (1968). Modification of classroom behavior of institutionalized female adolescent offenders. Beh. Res. and Therapy, 6, 343-353.

Meltzoff, J. and Kornreich, M. (1970). Research in Psychotherapy. New York: Atherton Press.

Mendoza, S., Good, T. and Brophy, J. (1972). Who Talks in Junior High Classrooms? Report Series No. 68, Research & Development Center for Teacher Education. Univ. of Texas.

Mercatoris, M. and Craighead, W.E. (1974). Effects of non-participant observation on teacher and pupil classroom behavior. J. Educ. Psychol., 66, 512-519.

Merrett, F. (1981). Studies in behaviour modification in British educational settings. In: K. Wheldall (Ed.): The Behaviourist in the Classroom. Birmingham: Educational Review Offset Publications No. 1.

Meyer, H., Borgatta, E. and Jones, W. (1965). Girls at Vocational High. New York: Russell Sage Foundation.

Meyers, C.E., Atwell, A.A. and Orpet, R.D. (1968). Prediction of fifth-grade achievement from kindergarten rest and rating data. Educ. & Psychological Measurement, 28, 457-463.

Miller, L.C., Barrett, C.L., Hampe, E. and Noble, H. (1972). Comparison of reciprocal inhibition, psychotherapy and waiting-list control for phobic children. J. Abnorm. Psychol., 79, 269-279.

Miller, W.B. (1962). The impact of a total community delinquency control project. Social Problems, Fall, 168-191.

Ministry of Education (1955). Report of the Committee on Maladjusted Children (the Underwood Report). London: H.M.S.O.

Mintz, J., Auerbach, A.H., Luborsky, L. and Johnson, M. (1973). Patient's, therapist's and observer's views of psychotherapy: a 'Rashoman' experience or a reasonable consensus. Brit. J. Med. Psychol., 46, 83-89.

Miron, C.D. (1971). Behavior modification: guide for teachers. Baltimore, Maryland: Behavioral Information & Technology.

Mischel, W. (1968). Personality and Assessment. Wiley: New York.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. Psychol. Bull., 80, 252-283.

Moore, T. (1974). Techniques of individual assessment. In: M. Chazan, T. Moore, P. Williams and J. Wright (Eds.): The Practice of Educational Psychology. London: Longman, pp 116-147.

Moreno, J.L. (1953). Who Shall Survive? Foundations of Sociometry, Group Psychotherapy and Psychodrama. New York: Beacon House.

Moss, G. and Childs, J. (1981). In-service training for teachers in behavioural psychology: problems of implementation. In: K. Wheldall (Ed.): The Behaviourist in the Classroom. Birmingham: Educational Review Offset Publications No. 1.

Moulin, E.K. (1970). The effects of client-centred counseling using play media on the intelligence achievement and psycholinguistic ability of underachieving primary school children. Elem. School Guid. Counsel., 5, 85-98.

Mulligan, G. and Douglas, J.W.B. (1963). Delinquency and symptoms of maladjustment: the findings of a longitudinal study. Proc. Roy. Soc. Med., 56, 1083-1086.

Mullin, L. (1979). Social factors and child management within schools. M.A. Thesis: University of Newcastle-upon-Tyne (Unpublished).

Murray, E.J. and Jacobson, L.I. (1978). Cognition and learning in traditional and behavioral psychotherapy. in: S.L. Garfield and A.E. Bergin (Eds.): Handbook of Psychotherapy and Behavior Change: An Empirical Analysis (2nd ed.). New York: Wiley, pp 661-687.

Nash, R. (1976). Teacher Expectations and Pupil Learning. London: Routledge and Kegan Paul.

Neligan, G.A., Prudham, D and Steiner, H. (1974). The Formative Years: Birth, Family and Development in Newcastle-upon-Tyne. Oxford: Oxford Univ. Press.

Nelson, R.O. (1974). An expanded scope for behavior modification in school settings. J. Sch. Psychol., 12, 276-287.

Nelson, W.J. and Birkimer, J.C. (1978). Role of self instruction and self reinforcement in the modification of impulsivity. J. Consult. Clin. Psychol., 46, 183.

Nicholson, J.N. and Gray, J.A. (1972). Peak shift, behavioural contrast and stimulus generalization as related to personality and development in children. Brit. J. Psychol., 63, 47-62.

Nietzel, M.T., Winett, R.A., Macdonald, M.L., and Davidson, W.S. (1977). Behavioral Approaches to Community Psychology. New York: Pergamon.

Northway, M.A. (1952). A Primer of Sociometry. Canada: University of Toronto Press.

Obler, M. and Terwillinger, R. (1970). Pilot study on the effectiveness of systematic desensitization with neurologically impaired children with phobic disorders. J. Consult. Clin. Psychol., 34, 314-318.

O'Connor, R.D. (1969). Modification of social withdrawal through symbolic modeling. J. Appl. Beh. Anal., 2, 15-27.

O'Connor, R.D. (1972). Relative efficacy of modeling, shaping and the combined procedures for modification of social withdrawal. J. Abnorm. Psychol., 79, 327-334.

O'Dell, S. (1974). Training parents in behavior modification: a review. Psychological Bulletin, 81, 418-433.

O'Leary, K.D. (1978). The operant and social psychology of token systems. In: A.C. Catania and T.A. Brigham (Eds.): Handbook of Applied Behavior Analysis: Social and Instructional Processes. New York: Halsted Press.

O'Leary, K.D. and Becker, W.C. (1967). Behavior modification of an adjustment class: a token reinforcement program. Exceptional Children, 33, 637-642.

O'Leary, K.D., Becker, W.C., Evans, M.B. and Saudargas, R.A. (1969). A token reinforcement program in a public school: a replication and systematic analysis. J. Appl. Beh. Anal., 2, 1, 3-13.

O'Leary, K.D. and Johnson, S.B. (1979). Psychological assessment. In: H.C. Quay and J.S. Werry (Eds.): Psychopathological Disorders of Childhood (2nd Ed.) New York: Wiley, pp 210-246.

O'Leary, K.D., Kaufman, K.F., Kass, R.E. and Drabman, R.S. (1970). The effects of loud and soft reprimands on the behavior of disruptive students. Exceptional Children, 37, 145-155.

O'Leary, K.D. and Kent, R. (1973). Behavior modification for social action: research tactics and problems. In: L.A. Hamerlynch, L.C. Handy and E.J. Mash (Eds.): Behavior Change: Methodology, Concepts and Practice. Champaign, Illinois: Research Press pp 69-96.

O'Leary, K.D., Kent, R.N., and Kanowitz, J. (1975). Shaping data collection congruent with experimental hypotheses. J. Appl. Beh. Anal., 8, 43-51.

O'Leary, K.D. and O'Leary, S.G. (1972). Classroom Management: The Successful Use of Behavior Modification. New York: Pergamon.

O'Leary, K.D. and O'Leary, S.G. (Eds.) (1977). Classroom Management: The Successful Use of Behavior Modification 2nd edition. New York: Pergamon.

O'Leary, K.D., Turkewitz, H. and Taffel, S.J. (1973). Parent and therapist evaluation of behavior therapy in a child psychological clinic. J. Consult. Clin. Psychol., 41, 279-283.

O'Leary, S.G. and Dubey, D.R. (1979). Applications of self-control procedures by children: a review. J. Appl. Beh. Anal., 12, 449-465.

O'Leary, S.G. and O'Leary, K.D. (1976). Behavior modification in the school. In: H. Leitenberg (Ed.): Handbook of Behavior Modification. Eaglewood Cliffs, New Jersey: Prentice Hall.

O'Leary, S.G. and Schneider, M.R. (1977). Special class placement for conduct problem children. Exceptional Children, 44, 24-30.

Parke, R.D. and Deur, J.L. (1972). Schedule of punishment and inhibition of aggression in children. Developmental Psychology, 7, 266-269.

Patterson, G.R. (1965). An application of conditioning techniques to the control of a hyperactive child. In: L.P. Ullman and L. Krasnov (Eds.): Case Studies in Behavior Modification. New York: Holt, Rinehart & Winston, pp 370-375.

Patterson, G.R. (1974). Interventions for boys with conduct problems: multiple settings, treatments and criteria. J. Consult. Clin. Psychol., 42, 471-480.

Patterson, G.R., Cobb, J.A. and Ray, R.S. (1973). A social engineering technology for retraining the families of aggressive boys. In: H.E. Adams and I.P. Unikel (Eds.): Issues and Trends in Behavior Therapy. Springfield, Illinois: Charles C. Thomas.

Patterson, G.R., Shaw, D.A. and Ebner, M.J. (1969). Teachers, peers and parents as agents of change in the classroom. In: F.A. Benson (Ed.): Modifying Deviant Social Behaviors in Various Classroom Settings. Eugene, Oregon: Univ. of Oregon.

Peterson, D.R. (1961). Behavior problems of middle childhood. J. Consult. Psychol., 25, 205-209.

Peterson, L., Homer, A.L. and Wonderlich, S.A. (1982). The integrity of independent variables in behavior analysis. J. Appl. Beh. Anal., 15, 477-492.

Peterson, R.F., Cox, M.A. and Bijou, S.W. (1971). Training children to work productively in classroom groups. Exceptional Children, 37, 491-500.

Petrie, I.R.J. (1962). Residential treatment of maladjusted children: a study of some factors related to progress in adjustment.

Pinkston, E.M., Reese, N.M., Leblanc, J.M. and Baer, D.M. (1973) Independent control of a preschool child's aggression and peer interaction by contingent teacher attention. J. Appl. Beh. Anal., 6, 115-124.

Platt, J.J., Spivack, G. and Swift, M.S. (1975). Interpersonal problem-solving group therapy. Philadelphia: Halinemann Medical College & Hospital.

Power, M.J., Benn, R.T. and Morris, J.N. (1972). Neighbourhood, school and juveniles before the courts. Brit. J. Criminol., April 111-132.

Powers, E. and Witmer, H.L. (1951). An Experiment in the Prevention of Delinquency. New York: Columbia Univ. Press.

Pring, R. (1981). Behaviour moficiation: some reservations. In: Behaviour Modification in Education Perspectives 5. Univ. of Exeter, pp 61-68.

Quay, H.C. (1979). Classification. In: H.C. Quay and J.S. Wern (Eds.): Psychopathological Disorders of Childhood (2nd edition). New York: Wiley, pp 1-42.

Quicke, J.C. (1975). Behaviourism and education. Association of Educational Psychologists Journal, 3, 8-16.

Rachman, S. and Wilson, G.T. (1980). The Effects of Psychological Therapy. Oxford: Pergamon Press.

Randolph, D.L. and Hardage, N.C. (1973). Behavioral consultation and group counselling with potential dropouts. Elementary School Guidance and Counseling, 7, 204-209.

Reason, P. and Rowan, J. (1981). Human Inquiry: a Sourcebook of New Paradigm Research. Chichester: John Wiley.

Redl, F. (1959). Strategy and techniques of the life space interview. Amer. J. Orthopsychiatry, 29, 1-18.

Registrar-General's Office. (1951), Classification of Occupations. London: H.M.S.O.

Reid, J.B. (1970). Reliability assessment of observation data: a possible methodological problem. Child Dev., 41, 1143-1150.

Reid, J.B. and Patterson, G.R. (1976). Follow up analyses of a behavioural treatment program for boys with conduct problems: a reply to Kent. J. Consult. Clin, Pscyhol.,

Reid, W.J. and Epstein, L. (1972). Task-Centred Casework. New York: Columbia University Press.

Reid, W.J. and Shyne, A.W. (1969). Brief and Extended Casework. New York: Columbia University Press.

Repucci, N.D. (1977). Implementation issues for the behavior modifier as institutional change agent. Behavior Therapy, 8, 594-605.

Repucci, N.D. and Saunders, J.T. (1974). Social psychology of behavior modification: problems of implementation in natural settings. American Psychologist, 29, 649-660.

Reynolds, D. (1976). The delinquent school. In: M. Hammersley and P. Woods (Eds.): The Process of Schooling. London: Routledge & Kegan Paul, pp 217-230.

Reynolds, G.S. (1975). A Primer of Operant Conditioning (Revised Edition). Glenview, Illinois: Scott, Foresman & Co,

Rachman, N. (1964). The prevalence of psychiatric disturbance in a hospital school for epileptics. Unpublished D.P.M. dissertation. Univ. of London.

Ringer, U.M.J. (1973). The use of a 'token helper' in the managment of classroom behavior problems and in teacher training. J. Appl. Beh. Anal., 6, 671-677.

Ritter, B. (1968). The group desensitization of children's snake phobias using vicarious and contact desensitization procedures. Beh. Res. & Ther., 6, 1-6.

Robins, L.N. (1972). Follow-up studies of behavior disorders in children. In: H.C. Quay and J. S. Werry (Eds.): Psychopathological Disorders of Childhood. New York: Wiley.

Robins, L.N. (1979). Follow-up studies. In: H.C. Quay and J.S. Werry (Eds.): Psychopathological Disorders of Childhood (2nd ed.) New York: Wiley, pp 483-514

Roedell, W.C., Slaby, R.G. and Robinson, H.B. (1977). Social development in young children. Monterey, California: Brooks/Cole.

Roff, M., Sells, S.B. and Golden, M.M. (1972). Social Adjustment and Personality Development in Children. Minneapolis: Univ. of Minnesota Press.

Rogers, C.R. (1952). Client-Centred Therapy. Boston: Houghton-Mifflin.

Rogers, C.R. (1959). A theory of therapy, personality and interpersonal relationships as developed in the client-centred framework. In: S. Koch (Ed.): Psychology: A Study of a Science. Vol. 3. New York: McGraw-Hill.

Rollins, H.A., McCandless, B.R., Thompson, M. and Brassell, W.R. (1974). Project Success Environment: an extended application of contingency management in inner city schools. J. Educ. Psychol., 66, 167-178.

Rollins, H.A. and Thompson, M. (1978). Implementation and operation of a contingency management program by the elementary school principal. Amer. Educ. Res. J.,15, 325-330.

Romanczyk, R.G., Kent, R.N., Diament, C. and O'Leary, K.D. (1973). Measuring the reliability of observational data: a reactive process. J. Appl. Beh. Anal., 6, 175-184.

Rose, G. and Marshall, A.M. (1974). Counselling and Social Work. Chichester: John Wiley.

Rosen, A. and Proctor, E.K. (1981). Distinctions between treatment outcomes and their implications for treatment evaluation. J. Consult. Clin. Psychol., 49, 418-425.

Rosenshine, B. (1970). Evaluation of classroom instruction. Review of Educational Research, 40, 279-300.

Rosenthal, A.J. and Levine, S.V. (1971). Brief psychotherapy with children: process of therapy. Amer. J. Psychiat., 128, 141-146.

Rosenthal, R. and Jacobson, L. (1968). Pygmalion in the Classroom. New York: Holt, Rinehart & Winston.

Ross, A.O. (1976). Psychological Aspects of Learning Disabilities and Reading Disorders. New York: McGraw-Hill.

Ross, A.O. (1981). Child Behavior Therapy. New York: Wiley.

Rosser, E. and Harre, R. (1976). The meaning of 'trouble'. In: M. Hammersley and P. Woods (Eds.): The Process of Schooling: A Sociological Reader. London: Routledge & Kegan Paul, pp 171-178.

Rothkopf, E.Z. (1968). Two scientific approaches to the management of instruction. In: R.M. Gagne and W.J. Gephart (Eds.): Learning Research and School Subjects. Itasca, Illinois: Peacock, pp 107-132.

Rotter, J.B., Liverant, S. and Crowne, D.P. (1961). The growth and extinction of expectancies in chance controlled and skilled tasks. J. Psychol., 52, 161-177.

Rowe, M. (1969). Science, silence and sanctions. Science and Children, 6, 11-13.

Rubin, R.A. and Balow, B.E. (1977). A longitudinal survey of school behavior problems Interim Report 25. U.S.O.E. Project 6-1176. Minneapolis: Univ. of Minnesota.

Rule, S.A. (1972). Comparison of three different types of feedback on teachers' performance. In: G. Semb (Ed.): Behavior Analysis and Education. Lawrence: Univ. of Kansas.

Rutter, M. (1965). Classification and categorization in child psychiatry. J. Child. Psychol. Psychiat., 6, 71-83.

Rutter, M. (1967). A children's behaviour questionnaire for completion by teachers: preliminary findings. J. Child Psychol. & Psychiat., 8, 1-11.

Rutter, M. (1982). Psychological therapies in child psychiatry: issues and prospects. Psychological medicine, 12, 723-740.

Rutter, M., Cox, A., Tupling, C., Berger, M. and Yule, W. (1975). Attainment and adjustment in two geographical areas: I. The prevalence of psychiatric disorder. Brit. J. Psychiat., 126, 493-509.

Rutter, M. and Graham, P. (1968). The reliability and validity of the psychiatric assessment of the child: I. Interview with the child. Brit. J. Psychiat., 114, 563-579.

Rutter, M., Graham, P., Chadwick, O, and Yule, W. (1976). Adolescent turmoil: fact or fiction? J. Child Psychol. Psychiat., 17, 35-56.

Rutter, M., Maughan, B., Mortimore, P. and Ouston, J. (1979). Fifteen Thousand Hours. london: Open Books.

Rutter, M., Tizard, J. and Whitmore, K. (1970). Education, Health and Behaviour. London: Longman.

Rutter, M. and Yule, W. (1977). Reading difficulties. In: M. Rutter and L. Hersov (Eds.): Child Psychiatry: Modern Approaches. London: Blackwell, pp 556-580.

Safer, D.J., Heaton, R.C. and Parker, F.C. (1981). A behavioral program for disruptive junior high school students: results and follow-up. J. Abnorm. Child Psychol., 9, 483-494.

Sajwaj, T. and Dillon, A. (1977). Complexities of an 'elementary' behavior modification procedure: differential adult attention used for children's behavior disorders. In: B.C. Etzel, J.M. Leblanc and D.M. Baer (Eds.): New Developments in Behavioral Research. Hillsdale, New Jersey: Lawrence Erlbaum. pp 303-317.

Samph, T. (1976). Observer effects on teacher verbal classroom behavior. J. Educ. Psychol., 68, 736-741.

Samuels,B.J. and Turnure, J.E. (1974). Attention and reading achievement in first grade boys and girls. J. Educ. Psychol., 66, 29-32.

Santogrossi, D.A., O'Leary, K.D., Romanczyk, R.G. and Kauffman, K.F. (1973). Self evaluation by adolescents in a psychiatric hospital school token program. J. Appl. Beh. Analysis, 6, 277-287.

Sarason, S.B. (1971). The Culture of the School and the Problem of Change. Boston: Allyn & Bacon.

Sargent, H.D. (1960). Methodological problems of follow-up studies in psychotherapy research. Amer. J. Orthopsychiat., 30, 495-506.

Saudargas, R.A. (1972). Setting criterion rates of teacher praise. The effects of video tape feedback in a behavior analysis follow-through classroom. In: G. Semb (Ed.): Behavior Analysis and Education. Lawrence: Univ. of Kansas, 1972, pp 253-261.

Schmuck, R.A. (1968). Helping teachers improve classroom group process. J. Appl. Beh. Sci., 4, 401-435.

Schnell, J.F. (1974). A brief report on invalidity of parent evaluations of behavior change. J. Appl. Beh. Anal., 7, 341-343.

Schultz, E.W., Hirshoren, A., Manton, A. and Henderson, R. (1971). Special education for the emotionally disturbed. Exceptional Children, 38, 313-320.

Schutte, R. C. and Hopkins, B.L. (1970). The effects of teacher attention on following instructions in a kindergarten class. J. Appl. Beh. Analysis, 3, 117-122.

Schwartz, S. and Johnson, J.H. (1981). Psychopathology of Childhood: A Clinical-Experimental Approach. New York: Pergamon.

Schwieso, J. and Hastings, N. (1981). The role of theory in the teaching of behaviour modification to teachers. In: K. Wheldall (Ed.): The Behaviourist in the Classroom. Birmingham: Educational Review Offset Publications No. 1.

Scott, W.A. and Wertheimer (1962). Introduction to Psychological Research. New York: Wiley.

Seeman, J., Barry, E. and Ellinwood, C. (1964). Interpersonal assessment of play therapy outcome. Psychotherapy: Theory, Research, Practice, 1, 64-66.

Sells,S.B. and Roff, M. (1967). Peer acceptance-rejection and personality development. Final Report, Project No. OE5-0417. U.S. Dept. of Health, Education and Welfare.

Semler, I.J. (1960). Relationships among several measures of pupil adjustment. J. Educ. Psychol., 51, 60-64.

Shapiro, D.A. (1980). Science and psychotherapy: the state of the art. Brit. J. Med. Psychol., 53, 1-10.

Shaw, H. (1952). A study of popular and unpopular children. Unpublished diploma: Univ. of Birmingham.

Shepherd, M., Oppenheim, A. and Mitchell, S. (1966). Childhood behaviour disorders and the child guidance clinic: an epidemiological study. J. Child Psychol. and Psychiat., 7,39-52.

Shepherd, M., Oppenheim, A.N. and Mitchell, S. (1971). Childhood Behavior and Mental Health. London: University Press.

Sherif, M. and Sherif, C.W. (1969). Social Psychology. New York: Harper & Row.

Sherman, J.A. and Bushell, D. (1974). Behavior modification as an educational technique. In: F.D. Horowitz (Ed.): Review of Child Development Research (Vol. 4). Chicago: University of Chicago Press.

Shipman, M. (1981). The Limitations of Social Research (2nd edition). London: Longman.

Shouksmith, G. and Taylor, J.W. (1964). The effect of counselling on the achievement of high-ability pupils. Brit. J. Educ. Psychol., 34, 51-57.

Shuller, D.Y. and McNamara, J.R. (1976). Expectancy factors in behavioral observation. Beh. Ther., 7, 519-527.

Shure, M.B. and Spivack, G. (1980). Interpersonal problem-solving as a mediator of behavioral adjustment in preschool and kindergarten children. J. Appl. Devel. Psychol., 1, 29-44.

Sidman, M. (1960). Tactics of Scientific Research: Evaluating Experimental Data in Psychology. New York: Basic Books.

Siegel, L.J., Dragovich, S.L. and Marholin,D. (1976). The effects of biasing information on behavioral observations and rating scales. J. Abnorm. Child Psychol., 4, 221-233.

Skaberne, B., Blejec, M., Skalev, V. and Vodopivec, K. (1965). Criminal prevention and elementary school children. Rev. Krim., 16, 8-14,.

Skinner, B.F. (1938). The Behavior of Organisms. New York: Appleton-Century-Crofts.

Skinner, B.F. (1953). Science and Human Behavior. New York: Macmillan.

Skinner, B.F. (1972). Beyond Freedom and Dignity. London: Cape.

Skinner, B.F. (1974). About Behaviorism. London: Jonathan Cape.

Sloane, R.B., Staples, F.R., Cristol, A.H., Yorkston, N.J. and Whipple, K. (1975). Short Term Analytically Oriented Psychotherapy Versus Behavior Therapy. Cambridge, Mass.: Harvard Univ. Press.

Smith, M.L. and Glass, G.V. (1977). Meta-analysis of psychotherapy outcome studies. American Psychologist, 32, 752-760.

Spielberger, C.D. and De Nike, L.D. (1966). Descriptive behaviorism versus cognitive theory in verbal operant conditioning. Psychol. Rev., 73, 306-326.

Spivack, G., Platt, J.J. and Shure, M.B. (1976). The Problem-Solving Approach to Adjustment. San Francisco: Jossey-Bass.

Spivack, G. and Shure, M. (1974). Social Adjustment of Young Children. San Francisco: Jossey-Bass.

Spivack, G. and Swift, M. (1966). The Devereux Elementary School Behavior Rating Scales: a study in the nature and organization of achievement related disturbed classroom behavior. J. Spec. Educ., 1, 71-91.

Spivack, G. and Swift, M. (1967). Devereux Elementary School Behavior Rating Scale Manual. Devon, Pennsylvania: The Devereux Foundation.

Staats, A.W. (1968). Learning, Language and Cognition. New York: Holt, Rinehart & Winston.

Staats, A.W. (1973). Behavior analysis and token reinforcement in educational behavior modification and curriculum research. In: C.E. Thoresen (Ed.): Behavior Modification in Education. Chicaco: Univ. of Chicago Press.

Staats, A.W., Minke, K.A., Finley, J.R., Wolf, M. and Brooks, L.O. (1964). A reinforcer system and experimental procedure for the laboratory study of reading acquisition. Child Development, 35, 209-231.

Stallings, J. (1975). Implementation and child effects of teaching practices in Follow Through classrooms. Monogr. Soc. Res. Child Dev., 40.

Stallings, J.A. and Kaskowitz, D.H. (1974). Follow Through Classroom Observation Evaluation. Menlo Park, California: Stanford Research Institute.

Stein, T.J. (1975). Some ethical considerations of short-term workshops in the principles and methods of behavior modification. J. Appl. Beh. Anal., 8, 113-115.

Stokes, T.F. and Baer, D.M. (1977). An implicit technology of generalization. J. Appl. Beh. Anal., 10, 349-367.

Stoltz, S.B. (1978). Ethical Issues in Behavior Modification. New York: Jossey-Bass.

Stott, D.H. (1963). The Social Adjustment of Children. London: Univ. of London Press.

Strain, P., Cooke, T. and Apolloni, T. (1976). Teaching exceptional children: assessing and modifying social behavior. New York: Academic Press.

Strain, P.S. and Shores, R.E. (1977). Social reciprocity: a review of research and educational implications. Exceptional Children, 43, 526-530.

Strickler, E. (1965). Educational group counseling within a remedial reading program. Diss. Abs. 25, 5129.

Stromer, R. (1977). Remediating academic deficiencies in learning disabled children. Exceptional Children, 43, 432-440.

Strupp, H. and Hadley, S.W. (1977). A tripartite model of mental health and therapeutic outcomes. American Psychologist, 32, 187-196.

Sulzer, B., Hunt, S., Ashby, E., Koniarski, C. and Krams, M. (1971). Increasing rate and percentage correct in reading and spelling in a 5th grade public school class of slow readers by means of a token system. In: E. Ramp and B.L. Hopkins (Eds.): A New Direction for Education: Behavior Analysis. Lawrence: The University of Kansas.

Surrat, P.R., Ulrich, R. and Hawkins, R.P. (1969). An elementary student as a behavioural engineer. J. Appl. Beh. Anal., 2, 85-92.

Swift, M.S. and Spivack, G. (1968). The assessment of achievement and related classroom behavior. J. Spec. Educ., 2, 137-153.

Swift, M.S. and Spivack, G. (1969). Achievement related classroom behavior of secondary school normal and disturbed students. Exceptional Children, 36, 677-684.

Swift, M., Spivack, G., Danset, A., Danset-Leger, J. and Winnykamen, F. (1972). Patterns of classroom behavioral diagnostic system for French and American Elementary Schools. Paper presented at the International Conference of Applied Psychology. Liege, Belgium.

Tait, C.D. and Hodges, E.F. (1962). Delinquents, Their Families and the Community. Springfield, Illinois: Charles C. Thomas.

Taplin, P.S. and Reid, J.B. (1973). Effects of instructional set and experimental influences on observer reliability. Child Dev., 44, 547-554.

Thomas, D.R. (1972). Self-monitoring as a technique for modifying teacher behaviors. Unpublished doctoral dissertation: Univ. of Illinois.

Thomas, D.R., Becker, W.C. and Armstrong, M. (1968). Production and elimination of disruptive classroom behavior by systematically varying teacher's behavior. J. Appl. Beh. Analysis, 1, 35-45.

Thomas, J.D. and Adams, M.A. (1971). Problems in teacher use of selected behavior modification techniques in the classroom. New Zealand J. Educ. Studies, 6, 151-165.

Thomas, J.D., Presland, I.E., Grant, M.D. and Glynn, T.L. (1978). Natural rates of teacher approval and disapproval in grade seven classrooms. Journal of Applied Behavior Analysis, 11, 91-94.

Thombs, M.R. and Muro, J.J. (1973). Group counseling and the sociometric status of second grade children. Elementary School Guidance and Counseling, 7, 194-197.

Tindall, R.H. (1955). Relationships among indices of adjustment status. Educ. Psychol. Measmt., 15, 152-162.

Tindall, R.H. (1959). Relationships among measures of adjustment. In: L. Gorlow and W. Katkovsky (Eds.): Readings in the Psychology of Adjustment. New York: McGraw Hill.

Tomlinson, J.R. (1972). Implementing behavior modification programs with limited consultation time. J. School Psychology, 10, 379-386.

Tramontana, M.G. (1980). Critical review of research on psychotherapy outcome with adolescents: 1967-1977. Psychol. Bull., 88, 429-450.

Trent, R.D. (1957). The relationship of anxiety to popularity and rejection among institutionalized delinquent boys. Child Dev., 28, 379-400.

Truax, C.B. (1968). Therapist interpersonal reinforcement of client self-exploration and therapeutic outcome in group psychotherapy. J. Counseling Psychology, 15, 225-231.

Truax, C.B. and Mitchell, K.M. (1971). Research on certain therapist interpersonal skills in relation to process and outcome. In A.E. Bergin and S.L. Garfield (Eds.): Handbook of Psychotherapy and Behavior Change: An Empirical Analysis. New York: Wiley.

Tsoi, M.M. and Nicholson, J.N. (1982). A factor analytic study of teachers' ratings of the personality of school children. Personality and Individual Differences, 3, 53-63.

Turkewitz, H., O'Leary, K.D. and Ironsmith, M. (1975). Generalization and maintenance of appropriate behavior through self-control. J. Consult. Clin. Psychol., 43, 577-583.

Twardosz, S. and Sajwaj, T. (1972). Multiple effects of a procedure to increase sitting in a hyperactive, retarded boy. J. Appl. Beh. Analysis, 5, 73-78.

Tyrer, P. and Tyrer, S. (1974). School refusal, truancy and adult neurotic illness. Psychological Medicine, 4, 416-421.

Uger, C. (1938). The relationship of teachers' attitudes to children's problem behaviour. School & Society, 47, 246-248.

Ullman, C.A. (1952). Identification of maladjusted school children. Public Health Monograph, No. 7.

Van Houten, R. and Sullivan, K. (1975). Effects of an audio cueing system on the rate of teacher praise. J. Appl. Beh. Anal., 8, 197-201.

Wahler, R.G. (1969). Setting generality: some specific and general effects of child behavior therapy. J. Appl. Beh. Analysis, 2, 239-246.

Walker, H.M. and Buckley,N. (1968). The use of positive reinforcement in conditioning attending behaviour. J. Appl. Beh. Analysis, 1, 245-252.

Walker, H.M. and Buckley, N.K. (1972). Programming generalization and maintenance of treatment effects across time and across settings. J. Appl. Beh. Analysis, 5, 209-224.

Walker, H.M., Greenwood, G.R., Hops, H. and Todd, N. (1979). Differential effects of reinforcing topographic components of social interaction. Behavior Modification, 3, 291-321.

Walker, H.M. and Hops, H. (1973). The use of group and individual reinforcement contingencies in the modification of social withdrawal. In: L.A. Hamerlynck, L.C. Handy and E.J. Mash (Eds.): Behavior Change: Methodology, Concepts and Practice. Champaign, Illinois: Research Press.

Walker, H.M. and Hops, H. (1976). Increasing academic achievement by reinforcing direct academic performance and/or facilitative nonacademic responses. J. Educ. Psychology, 68, 218-225.

Walker, H.M., Hops, H. and Johnson, S.M. (1975). Generalization and maintenance of classroom treatment effects. Behavior Therapy, 6, 188-200.

Walker, H.M., Mattson, R.H. and Buckley, N.K. (1971). The functional analysis of behavior within an experimental class setting. In: W.C. Becker (Ed.): An Empirical Basis for Change in Education. Chicago: S.R.A. pp 236-263.

Walter, H. and Gilmore, S.K. (1973). Placebo versus social learning effects in parent training procedures designed to alter the behaviours of aggressive boys. Beh. Res. & Ther., 4, 361-377.

Wanlass, R.L. and Prinz, R.J. (1982). Methodological issues in conceptualizing and treating childhood social isolation. Psychol. Bull., 92, 39-55.

Ward, J. (1971). Modification of deviant classroom behaviour. British Journal of Educational Psychology, 41, 3, 304-313.

Ward, M.H. and Baker, B.L. (1968). Reinforcement therapy in the classroom. J. Appl. Beh. Analysis, 1, 323-328.

Warren, W.A. (1965). A study of adolescent psychiatric in-patients and the outcome six or more years later. II The follow up study. J. Child Psychol. & Psychiat.,6, 141-160.

Wasik, B.H., Senn, K., Welch, R.H. and Cooper, B.R. (1969). Behavior modification with culturally deprived school children: two case studies. J. Appl. Beh. Anal., 2, 181-194.

Watson, G.A. (1933). A critical note on two attitude scales. Mental Hygiene, 17, 59-64.

West, D.J. (1967). The Young Offender. Harmondsworth: Penguin.

West, D.J. and Farrington, D.P. (1973). Who Becomes Delinquent? London: Heineman.

Wheldall, K. (1981a). The Behaviourist in the Classroom. Birmingham: Educational Review Offset Publications No. 1.

Wheldall, K. (1981b). Concluding comments: 'A before C' or the use of behavioural ecology in the classroom. In: K. Wheldall (Ed.): The Bheaviourist in the Classroom. Birmingham: Educational Review Offset Publications, No. 1 pp 121-130.

Wheldall, K. and Austin, R. (1980). Successful behaviour modification in the secondary school: a reply to McNamara and Harrop. B.P.S.: Division of Educational and Child Psychology. Occasional Papers, 4, 3-9.

Wheldall, K. and Merrett, F. (1982). The behavioural approach to teaching (BAT Project). Association for Child Psychology and Psychiatry News, No. 13, 2-7.

White, M.A. (1975). Natural rates of teacher approval. J. Appl. Beh. Anal., 8, 367-372.

Wickman, E.K. (1928). Children's Behaviour and Teachers' Attitudes. Oxford: Oxford Univ. Press.

Wildman, R.W. and Wildman, R.W. (1975). The generalization of behavior modification procedures: a review with special emphasis on classroom applications. Psychol. in the Schools, 12, 432-448.

Wiley, D.E. and Harnischfeger (1974). Explosion of a myth: quantity of schooling and exposure to instruction, major educational vehicles. Studies of Educative Processes, Report No. 8. University of Chicago.

Williams, J.L. (1973). Operant Learning: Procedures for Changing Behavior. Monterey: California: Brooks/Cole.

Wiltz, N.A. and Patterson, G.R. (1974). An evaluation of parent training procedures designed to alter inappropriate aggressive behavior of boys. Beh. Ther., 5, 515-521.

Winer, B.J. (1971). Statistical Principles in Experimental Design - 2nd edition. London: McGraw-Hill.

Winett, R.A. and Winkler, R.C. (1972). Current behavior modification in the classroom: be still, be quiet, be docile. J. Appl. Beh. Analysis, 5, 499-504.

Winett, R.A. and Roach, E.M. (1973). The effects of reinforcing academic performance on social behavior. Psychological Record, 23, 391-396.

Winkler, R.C. (1971). The relevance of economic theory and technology of token reinforcement systems. Beh. Res. & Ther., 9, 81-88.

Winkler, R.C., Teigland, J.J. and Munger, P.F. (1965). The effects of selected counselling and remedial techniques on underachieving elementary school students. J. Counseling Psychol., 12, 384-387.

Witmer, H.L. and Keller, J. (1942). Outgrowing childhood problems: a study of the value of child guidance treatment. Smith College Studies in Social Work, 13, 74-90.

Wodarski, J.S. and Filipczak, J. (1982). Behavioral intervention in public schools. II: Long-term follow-up. In: D.J. Safer (Ed.): School Programs for Disruptive Adolescents. Baltimore: University Park Press, pp 201-214.

Wolf, M.M., Giles, D.K. and Hall, V.R. (1968). Experiments with token reinforcement in a remedial classroom. Beh. Res. & Ther., 6, 51-64.

Wolf, M.M., Hanley, E.L., King,L.A., Lachowicz, J. and Giles, D.K. (1970). The timer-game: a variable interval contingency for the management of out-of-seat behavior. Exceptional Children, 37, 113-117.

Wood, F.H. and Zabel, R.H. (1978). Making sense of reports on the incidence of behaviour disorders/emotional disturbance in school aged populations. Psychology in the Schools, 15, 45-51.

Woods, P. (1976). Having a laugh: an antidote to schooling. In: M. Hammersley and P. Woods (Eds.): The Process of Schooling: A Sociological Reader. London: Routledge & Kegan Paul, pp 178-187.

Woods, P. (1980). Pupil Strategies: Explorations in the Sociology of the School. London: Groom-Helm.

Woods, P. (1980). Teacher Strategies: Explorations in the Sociology of the School. London: Groom-Helm.

Woolfolk, A.E., Woolfolk, R.L. and Wilson, G.T. (1977). A rose by any other name... labelling bias and attitudes towards behaviour modification. J. Counsell. Clin. Psychol., 45, 184-191.

Workman, E.A., Kindall, L.M. and Williams, R.L. (1980). The consultative merits of praise-ignore versus praise-reprimand instruction. J. School Psychology, 18, 373-380.

Wortman, P.M. (1975). Evaluation research: a psychological perspective. American Psychologist, 30, 562-575.

Wright, D.M., Moelis, I. and Pollack, L.J. (1976). The outcome of individual psychotherapy: increments at follow-up. J. Child Psychol. & Psychiat., 17, 275-285.

Yalom, I.D. (1975). The Theory and Practice of Group Psychotherapy. (2nd edition). New York: Basic Books.

Yeaton, W.H. (1982). A critique of the effectiveness of applied behaviour analysis research. Adv. Behav. Res. Ther., 4, 75-96.

Yellott, A.W., Liem, G.R. and Cowen, E.L. (1969). Relationships among measures of adjustment, sociometric status and achievement in third grades. Psychology in the Schools, 6, 315-321.

Yule, W. (1977). Behavioural approaches. In: M. Rutter and L. Hersov (Eds.): Child Psychiatry: Modern Approaches. Oxford: Blackwell, pp 923-947.

Zeiler, M.D. and Harzem, P. (Eds.) (1979). Reinforcement and the Organization of Behaviour. Chichester: John Wiley.

Zimmerman, E.H., and Zimmerman, J. (1962). The alteration of behavior in a special classroom situation. J. Experimental Analysis of Behavior, 5, 59-60.