

ACCEPTABILITY OF ALTERNATIVE CLASSROOM TREATMENT STRATEGIES
AND FACTORS AFFECTING TEACHER'S RATINGS

by

WILLIAM TREEN MCKEE

B.A., The University of British Columbia, 1972

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

Department of Educational Psychology and Special Education

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

December, 1984

© William Treen McKee, 1984

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

Department of Educational Psychology and Special Education

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date: December, 1984

Abstract

A screening sample of 107 elementary school teachers completed a 16-item multiple-choice measure of knowledge of behavioral principles. Two subject groups, each consisting of 32 regular classroom teachers, were randomly selected for high and low knowledge groups based on scores on the knowledge measure. High and low knowledge group teachers were given random assignment to one of two clinical case descriptions. Both descriptions were of a 9-year old boy who is presented as being very aggressive, and who does not follow adult instructions. The two cases differed only in the type of background attributed to each. One is described as coming to the regular classroom as a result of 'mainstreaming', having been previously in a special class setting. The other case description makes no reference to mainstreaming or to previous placement. Teachers in all conditions evaluated the acceptability of four alternative classroom treatment strategies for children: Medication, Time Out from Reinforcement, Reinforcement of Incompatible Behavior, and Positive Practice. Teachers in the high and low knowledge groups read the assigned case description, and then rated the acceptability of the treatments in a 4 x 4 replicated Latin-square design using treatment acceptability measures developed by Kazdin (1980a). Following treatment ratings, subjects completed a questionnaire which, in part, asked subjects to rank order the described treatments on the basis of likelihood of each being used by the subject doing

the rating. High knowledge group teachers rated treatments as more acceptable than did low knowledge group teachers on the primary dependent measure. Treatments were readily differentiated in terms of acceptability. Reinforcement was rated more acceptable than Time Out and Positive Practice, which did not differ from each other in terms of acceptability. Medication was rated lower in acceptability than the other treatments. Treatment acceptability ratings were not different for the two case descriptions. The results of the ranking of treatments on the Case and Treatment Questionnaire generally followed those of the rating procedure. The results of the analyses indicate that teachers do differentiate treatments in terms of acceptability using both the rating and ranking procedures. It is also apparent that the high knowledge group teachers tend to rate alternative classroom treatments as more acceptable than do the low knowledge group teachers. These findings are discussed, and implications for education and for future research are presented.

Table of Contents

Abstract	ii
List of Tables	vii
List of Figures	ix
Acknowledgement	x
I. DEVELOPMENT OF THE STUDY	1
A. INTRODUCTION	1
B. BACKGROUND OF THE PROBLEM	4
C. PROBLEM AREA	14
1. Knowledge of Behavioral Principles	17
2. Type of Case	20
3. Method of Assessing Acceptability	26
D. PROBLEM SUMMARY	28
E. PURPOSE OF THE STUDY	29
F. SIGNIFICANCE OF THE STUDY	30
II. METHODOLOGY	34
A. NATURE OF THE STUDY	34
B. INSTRUMENTATION	35
1. Demographic Questionnaire	36
2. Knowledge Measure	36
3. Treatment Acceptability Measures	43
4. Case and Treatment Questionnaire	49
C. DESIGN OF THE STUDY	53
1. Dependent Variables	53
2. Independent Variables	54
3. Experimental Design	55
D. PROCEDURE	58

1. Subjects	58
2. Assignment of Subjects	60
3. Data Collection	61
E. DATA PREPARATION AND ANALYSIS	62
1. Data Preparation	62
2. Analysis of Demographic Questionnaire	63
3. Psychometric Analysis	63
4. Analysis of Treatment Acceptability Ratings	64
5. Analysis of Case and Treatment Questionnaire Responses	70
F. CHAPTER SUMMARY	70
III. RESULTS	72
A. SUBJECTS	72
B. PSYCHOMETRIC ANALYSIS	74
1. Knowledge Measure	74
2. Treatment Acceptability Measures	74
C. ANALYSIS OF TREATMENT ACCEPTABILITY RATINGS	79
1. Treatment Evaluation Inventory	79
2. Semantic Differential	87
3. Summary of Treatment Acceptability Results ..	94
D. CASE AND TREATMENT QUESTIONNAIRE ANALYSIS	98
IV. SUMMARY AND CONCLUSIONS	101
A. SUMMARY OF METHODS AND RESULTS	101
1. Knowledge Group	103
2. Child Case	105
3. Treatment Type	108
B. IMPLICATIONS FOR EDUCATION	111

C. IMPLICATIONS FOR FUTURE RESEARCH	114
REFERENCES	117
APPENDIX A	133
APPENDIX B	137
APPENDIX C	144
APPENDIX D	147
APPENDIX E	148
APPENDIX F	150
APPENDIX G	153

List of Tables

Table 1 – KBPAC Pilot Study Results	42
Table 2 – Characteristics of Screening Sample	60
Table 3 – Subject Demographic Characteristics	73
Table 4 – KBPAC Item Analysis	75
Table 5 – KBPAC Test Analysis	76
Table 6 – TEI Test Analysis	77
Table 7 – Semantic Differential Test Analysis	78
Table 8 – Anova For TEI	80
Table 9 – Tukey (HSD) Comparisons Between Treatment Means on the TEI	82
Table 10 – Tukey (HSD) Comparisons Between Order Means on the TEI	83
Table 11 – ANOVA for SD Evaluative Dimension	88
Table 12 – Tukey (HSD) Comparisons Between Treatment Means on the SD Evaluative Dimension	90
Table 13 – Tukey (HSD) Comparisons Between Order Means on the SD Evaluative Dimension	91
Table 14 – Tukey (HSD) Comparisons Between Sequence Means on the SD Evaluative Dimension	91
Table 15 – ANOVA for SD Potency Dimension	93
Table 16 – Tukey (HSD) Comparisons Between Treatment Means on the SD Potency Dimension	94
Table 17 – ANOVA for SD Activity Dimension	95
Table 18 – Tukey (HSD) Comparisons Between Order Means on the SD Activity Dimension	96
Table 19 – Summary of Tukey (HSD) Comparisons Between Treatment Conditions	97
Table 20 – Summary of Tukey (HSD) Comparisons Among Orders of Presentation	97

Table 21 - Results of Friedman Two-Way Analysis of Variance on Ranking of Treatments	99
---	----

Table 22 - Nemenyi Pairwise Comparisons Between Mean Treatment Ranks	100
---	-----

List of Figures

Figure 1 – Layout of the Experimental Design	57
Figure 2 – LSR Effects--CxFxB Residuals: Reinforcement and Positive Practice for Two Knowledge Conditions	.84
Figure 3 – LSR Effects--CxDxB Residuals: Reinforcement and Medication for Two Case Descriptions	86

Acknowledgement

I wish both to acknowledge and to express my appreciation to a number of individuals for many contributions made during the course of this study.

First, I wish to express my sincere appreciation to my research supervisor, Dr. Barry Munro, who has provided constant support and thoughtful advice, and who has been a model for the pursuit of academic excellence throughout our association. I wish also to thank my research committee members, Dr. Todd Rogers and Dr. Robert McMahon for their guidance, constructive criticism and generous support throughout the project. The contributions of my committee are gratefully acknowledged.

Two additional Faculty members made vital contributions at crucial stages in the study and deserve a special note of thanks. To Drs. G. Jim Johnson and Robert Conry, a word of appreciation for their involvement.

An expression of thanks also is due to Drs. Alan Kazdin and Joe Witt for their encouragement (from afar) in pursuing this study, and for the provision of otherwise unavailable materials. A special thanks to Dr. Kazdin for allowing the use and amendment of research instruments and case and treatment descriptions.

My appreciation also goes to the Burnaby school teachers who gave their time and energies to participate in the study procedures, and to Dr. Blake Ford for his efforts to facilitate teacher participation in the project.

Finally, I give thanks to my wife Marie, and to my daughter Caitlin, both of whom have supported my efforts unfailingly and who have tolerated my absences and preoccupations over several years of part-time graduate study. Marie, particularly for her tireless efforts toward perfection, and her well timed encouragement, has made a sizable contribution to this study.

I. DEVELOPMENT OF THE STUDY

A. INTRODUCTION

The acceptability and credibility of treatment approaches developed from the behavioral psychology paradigm have long been at issue. In the relatively early days of behaviorism, Dr. Louis Berman (1927) suggested that:

. . . a new, a powerful religion is growing into maturity in the United States as a result of a new psychological movement. It calls itself Behaviorism . . . it may be assuredly predicted that laws will be passed and policemen paid to make the very name of Behaviorism anathema. (pp. 9-10)

The behaviorism to which Berman refers was, in a historical perspective, not in its maturity, but rather in its infancy. In his Presidential Address to the Association for Behavior Analysis (ABA) in 1980, Jack Michael stated that,

Prior to 1930 there wasn't much going on In the period from 1930 to 1938, Skinner [B. F. Skinner] put it all together. He managed, within that brief but fertile period, to come up with almost all of the essential methods, concepts, and functional relations of our field as we see it today. (p. 1)

Behaviorism has indeed persevered. Behavioral approaches to the study of human behavior have evolved and, in fact, become pervasive (see, for example, Benassi & Lanson, 1972; Martin, 1981; O'Leary, 1984; Wilson, 1982). In many ways, however, the predictions made early on by Dr. Berman (1927) have come to pass. Behaviorism, behavior modification, and a multitude of approaches developed out of the behavioral paradigm are the subject of enduring public interest and scrutiny. These continue to be the subject of

debate and attempts at resolution in several political arenas, as well as being the issue of litigation before the courts, and the subject of investigation and evaluation in a number of professional, academic, and special interest organizations. Farkas (1981), as well as others (Eysenck, 1970, 1971; Eysenck & Beech, 1971), has documented a multitude of criticisms of behavioral approaches. Numerous others have made presentations discussing these same issues (Bandura, 1975, 1977; Barnes, Fors & Decker, 1980; Berger, 1979; Bergin, 1980; Eber & O'Brien, 1982; Engelhardt & McCullough, 1982; Farnham-Diggory, 1981; Heldman, 1973; Holland, 1975; Kazdin, 1979, 1981b, 1981c, 1982; Kitchener, 1980a, 1980b, 1980c; Krasner, 1976; Marotz, 1983; Roos, 1974; Skinner, 1971; Spece, 1972; Szasz, 1978; Tyler, 1982; Wolpe, 1978).

Fifty years after Berman's text, Arnold Lazarus (1977) suggested that: ". . . behavior therapy and behavior modification have acquired a bad press. To receive funding many hospitals and community agencies have had to drop the label behavior from their program proposals" (p. 553). It would appear that there remain significant concerns with the issue of acceptability of behavioral approaches to treatment.

Although treatments from a variety of orientations are the subject of probing and scrutiny on many dimensions of acceptability, several authors have recently considered legal and ethical dimensions of acceptability related to

behavior modification (Christian, Clark, & Luke, 1981; Dipasquale, 1979; Feldman & Peay, 1982; Goldiamond, 1975a, 1975b; Innis, 1981; James & Allon, 1979; Wherry, 1983; Woods, 1982). Acceptability in the context of this literature relates to development, regulation, and maintenance of ethical and legal behavioral treatments. From another, though related perspective, Wolf (1978) introduced a dimension of acceptability to be conceptualized as social validity. He proposed that behavior analysts develop systems and measures for asking society whether they (behavior analysts) were accomplishing something of social importance. Wolf suggested that validation occur at three levels. These involve a demonstration of:

1. The social significance of the treatment goals.
2. The social appropriateness of the treatment procedures.
3. The social importance of the treatment effects.

Stephanie Stolz (1981) has suggested that technological and methodological innovations from the behavioral paradigm have experienced little widespread adoption by our society primarily because these innovations have not been accepted by government policy makers. Wolf (1978) expressed a similar sensitivity to issues of acceptability of treatments to participants:

. . . that if the participants don't like the treatment then they may avoid it, or run away, or complain loudly. And thus, society will be less likely to use our technology, no matter how potentially effective and efficient it might be. (p. 206)

Other authors indicate that treatment procedures with demonstrated effectiveness are not being used "because they are simply unacceptable to participants, care givers, or consumers" (Witt, Martens, & Elliot, 1984, p. 1). Treatments which are obviously and totally unacceptable are often easy to identify. However of more acute concern are treatments which have been demonstrated to be effective and useful but which gain limited acceptance and hence are not used. Finally, McMahon and Forehand (1983) in a review of consumer satisfaction literature in the behavioral treatment of children, refer to work by Kazdin (1980a, 1980b, 1981a) which they feel "has developed a compelling case for the assessment of treatment acceptability on the basis of legal and ethical issues and long-term effectiveness and generality of the procedures." (p. 214)

The issue of treatment acceptability has been established as a legitimate area of investigation equal in importance, perhaps, to that of consumer satisfaction with therapeutic interventions (Kiesler, 1983). However, in contrast to research in the latter area, the literature related to treatment acceptability is far less developed.

B. BACKGROUND OF THE PROBLEM

Few studies have been reported which investigate the reasons for the lack of acceptance of behavioral methods for the classroom treatment of children's problem behavior. This paucity exists despite the knowledge and public lamentation

of the 'under utilization' of these methods, and despite the overwhelming body of theoretical and empirical evidence in support of the effectiveness of such methods and procedures (Barlow, 1980; Boutilier, 1981; Foxx & Jones, 1978; Foxx & Shapiro, 1978; Giles, 1983; Kazdin & Hersen, 1980; Kazdin & Wilson, 1978b; Kent & O'Leary, 1976; Kerr & Lambert, 1982; Kirigin, Braukmann, Atwater, & Wolf, 1982; Molloy, 1980 ; Ollendick, Matson, Esveltd-Dawson, & Shapiro, 1980; Pevsner, 1982; Reese, Murphy, & Filipczak, 1981; Rosenbaum, O'Leary, & Jacob, 1975; Seaton & Aaron, 1978; Treiber & Lahey, 1983; Wilson, 1982).

Some investigators, notably Kazdin and Cole (1981), Witt, Moe, Gutkin, and Andrews (in press), and Woolfolk, Woolfolk, and Wilson (1977) have proceeded in the area of global evaluations of behavior modification and, in particular, the effect of the 'behavioral' label on evaluations of treatment. These and numerous other studies have been directed at ascertaining the public's conception of behavior modification, using a large variety of methods, settings and subjects. Several projects have focused attention on the attitudes held by undergraduate and graduate psychology students toward the use of behavior modification (e.g., Kazdin & Cole, 1981; Turkat, Harris, & Forehand, 1979; Woolfolk & Woolfolk, 1979; Woolfolk et al., 1977). At least one study (Dubno, Hillburn, Robinson, Sandler, Trani, & Weingarten, 1978) has investigated graduate business students' attitudes toward behavior

modification in business organizations. Another study by Barling and Wainstein (1979), using a methodology similar to Woolfolk and Woolfolk (1979), assessed the attitudes of undergraduate and graduate industrial psychology students. Several reports have presented results of investigations of teachers' attitudes toward the implementation of behavior modification in educational settings (e.g., Musgrove & Harms, 1975; Throll & Ryan, 1975; Ryan, 1976; Vane, 1972; Wheldall & Congreve, 1980a, 1980b). A further area of investigation is reported by Saunders and Reppucci (1978) and reference is made to an earlier study by Saunders (1975) in which superintendents of juvenile corrections institutions and elementary school principals were studied with respect to their attitudes toward behavior modification.

The results of this large variety of studies are inconclusive. While several authors suggest that attitudes toward behavior modification are in general negative (Burton, 1981; Kazdin & Cole, 1981; O'Leary, 1984; Saunders & Reppucci, 1978; Stainbrook & Green, 1982; Turkat & Forehand, 1980a; Turkat et al., 1979; Turkat & Feuerstein, 1978; Woolfolk & Woolfolk, 1979; Woolfolk et al., 1977), several others report more positive attitudes (Burkhart, Behles, & Stumphauzer, 1976; Frey, 1970; Hickey, 1977; Jeger & McClure, 1979; Ryan, 1976; Vane, 1972; Wheldall & Congreve, 1980a, 1980b. Further studies report almost neutral attitudes (Barling & Wainstein, 1979; Musgrove,

1974; Throll & Ryan, 1976). These contradictory findings would suggest that the question of attitude toward behavior modification has not been answered completely. While inconclusive in this regard these studies have provided valuable data furthering our understanding of the study of attitudes toward behavior modification.

There is some support, however, for the notion that this research suffers from some fundamental problems. As Kazdin and Cole (1981) suggest, the behavior modification under such considered scrutiny in these studies may not exist: "research typically discusses 'behavior modification' as a uniform method of treatment. In fact, several very different techniques are embraced by behavior modification and behavior therapy (terms usually used synonymously)" (p. 66). Many techniques, considered to be examples of behavior modification, may be evaluated differently not only because of actual differences in the techniques but also because of differences in problems and clients to which they might be applied. It could be argued that, in fact, individuals may be assessing behavior modification based on rather erroneous information. As reported by Turkat and Feuerstein (1978) and echoed by O'Leary (1984), media coverage of behavior modification is frequently negative, and often, incorrectly associated with such procedures as brain washing, psychosurgery, sensory deprivation, and water torture. Friedman (1975) suggests that the public view of a link between behavior modification and any number of

questionable therapies, as well as a confusion between behavior modification as an end product and as a specific procedure, has caused "all the abuses of the psychosurgeons or the psychopharmacologists to be laid at the doorstep . . . (p. 47) of behavior modification.

Kazdin and Cole (1981) suggest further that the direction for investigations of treatment acceptability should change from that of global acceptability to investigations of "the full gamut of procedures and the manner in which these might be investigated" (p. 67). What is suggested, then, is the study of particular procedures with consideration of the particular effect of such factors as the settings in which the procedure is to be applied, the clients and types of problems for which these procedures might be appropriate and the individuals who may be in a position to influence decisions regarding their application, or who may, in fact, be the subject of their application.

Initial efforts have begun to create approaches for the study of acceptability of particular treatment alternatives. Characteristic of these early studies is a great variety in emphases. One focus has been upon clinical or institutional settings (Kazdin, 1980a; Kazdin, 1980b; Kazdin, 1981). Other studies have emphasized student, parent and teacher satisfaction with particular behavioral approaches used in classroom projects (Besalel-Azrin, Azrin, & Armstrong, 1977; Kent & O'Leary, 1976), the emphasis here being both on the classroom setting for treatment and the use of a post hoc or

consumer evaluation of the treatment(s) employed. Witt and his colleagues (Witt, Elliot, & Martens, in press; Witt, Moe, Gutkin, & Andrews, in press; Witt et al., 1984) have approached the issue of the acceptability of treatment using classroom teachers as raters of treatments in several analogue studies. Norton, Austen, Allen, & Hilton (1983) report on a study comparing teacher and parent evaluations of the acceptability and effectiveness of several behavioral procedures for reducing children's disruptive behaviors. Foxx and Shapiro (1978) report using professionals and paraprofessionals, working with developmentally delayed children, as raters of treatments.

Considering these early studies of the acceptability of particular treatment alternatives, two major approaches can be identified. The first of these involves a post hoc evaluation of the application of a particular set of procedures, or consumer satisfaction measure, following experience with a particular approach to treatment. The consumer in this sort of evaluation might be conceptualized as a parent of a child in treatment, the teacher or therapist who used the approach, the child who has been the subject of treatment or other individuals involved in some way with the application of a particular mode or method of treatment. The second major approach involves individuals in the evaluation of procedures prior to any application of treatments. These evaluations of treatment acceptability are commonly based on fairly comprehensive descriptions of the

individual procedure being evaluated and would include information such as the types of cases and settings in which the procedure might be applied.

McMahon and Forehand (1983) have reviewed much of the literature dealing with consumer satisfaction in behavioral treatment of children. They reviewed studies which examine satisfaction with treatment outcome, therapists, treatment procedures, and teaching format, but suggest that the major portion of research in this area has focused on satisfaction with therapist and treatment outcome. These authors report further that the most pervasive use of consumer satisfaction measures is found in the parent training literature.

Other authors have expressed support for the inclusion of consumer satisfaction measures in the evaluation of treatments (Bornstein & Rychtarik, 1983; Garfield, 1983; Kiesler, 1983; Kirigin et al., 1982; Lebow, 1982). However, Scheirer (1978) has suggested caution in the interpretation and use of such consumer measures--a sentiment shared by others (Garfield, 1983; Kiesler, 1983). Guba (cited in Scheirer, 1978) provides the proposition that one significant problem with the measurement of client satisfaction is that "participants like social programs, evaluate them favorably, and think they are beneficial, irrespective of whether measureable behavioral changes take place toward stated program goals" (p. 55). McMahon and Forehand (1983) echoed this concern and concluded that: "While many investigators have included consumer

satisfaction measures in their treatment evaluations, there are serious methodological shortcomings in the development of these instruments as well as in their utilization as outcome measures" (p. 222), and recommend that adequate measures be developed before attention be given to absolute or comparative assessments of consumer satisfaction.

Witt and Elliot (in press) have reviewed much of the literature which has employed the second method of assessing acceptability of treatment alternatives. This report focused primarily on material relevant to the acceptability of classroom management strategies but reviewed the development and application of this method in other settings as well. A major conclusion of this review was that significant contributions to understanding the assessment of treatment acceptability have been made, even though research in this area is at a pioneering stage. Kiesler (1983) suggests additionally, that in the study of consumer satisfaction with treatment, researchers have not established the clients' general attitudes or expectation for treatment prior to entering, and cannot therefore suggest how much of post treatment evaluation relates to the treatment experience. This is a common methodological criticism of such research (Hargreaves & Atkisson, 1978; Lebow, 1982). Kiesler (1983) suggested the need for an addition to consumer satisfaction research of "the methodological equivalent of an attitude pretest" (p. 228). He suggests support for the 'pretest' approach employed by Kazdin (in press), and that such an

approach be used to study the public policy aspects of the acceptability of mental health services. Witt and Elliot (in press) also argue for the validity of the study of treatment acceptability and particularly this approach. They suggest that in the complex setting of school (for example) "it becomes important and desirable to expand the criteria by which various treatments are evaluated and to examine not only effectiveness but also the perceptions of individuals within a system concerning the application and effects of treatment" (p. 4-5). Witt and Elliot in describing future directions for research in this area suggest that empirical support is needed to establish the acceptability of various treatments, and that research will need to elucidate specific features of treatments (and presumably of those who rate the treatments) which contribute to evaluations of treatment acceptability. They suggest further that research focus on implementation of classroom interventions by teachers and on variables which may increase the probability that an intervention will be used. This appears to be a valuable direction for study since a major source for treatment selection and for implementation in the school setting is the classroom teacher of the child identified in need of some change. The major portion of reported studies of treatment acceptability to date have employed as raters undergraduate psychology students, not consumers or prospective consumers of the procedures being rated. More ideal raters would include classroom and special education

teachers, students in both regular and special education settings, those who consult directly with teachers, and others who directly influence the kinds of treatments which might be implemented--faculties of education and district inservice and special service personnel, for example.

Witt et al. (1984) and others (Baer & Bushell, 1981; Witt, Hannefin, & Martens, in press; Wolf, 1978) have expressed concern that many effective treatment procedures remain unused because they appear unacceptable to individuals who are in a position to effect decisions regarding selection of treatments. Kazdin (in press) suggests that acceptability of treatments has begun to receive attention because of the "increased participation of the courts, professional organizations, and institutional review committees in the legal and ethical issues raised by treatment" (p. 6). Despite the overwhelming volume of literature and empirical evidence which exists to support the effectiveness of most behavioral treatment methods, the link between the development of effective procedures and the widespread acceptance and implementation of these procedures has not been made. The problem may be one of building "better bridges from research to state hospitals [and schools]" (Stolz, 1981, p. 503), or one of better 'marketing' of effective strategies (Witt & Elliot, in press), or of changing the kinds of labels used (Kelly, 1950; Woolfolk et al., 1977). More realistically perhaps, the problem may require the development of a methodology for

ascertaining those variables relevant to evaluation such that procedures might be developed, training given, or information disseminated which will lead to better treatment being implemented in classrooms (as well as other settings).

C. PROBLEM AREA

Initial investigations of treatment acceptability have identified a host of factors which appear to affect evaluation ratings. These factors include: the effectiveness and appropriateness of the treatment in general, and for the particular type of client in question; the type of intervention; the method of assessing acceptability (e.g., unidimensional or multidimensional assessment [Lebow, 1982]); practical implementation issues such as time and resources, and who will perform the intervention; the theoretical orientation of the treatment; client or rater characteristics; and the language and labels attached to the treatment(s) under consideration.

A major portion of the treatment acceptability research presently available can be attributed to the work of Alan Kazdin and his colleagues at the School of Medicine, University of Pittsburgh. A particular contribution has been the development of a methodology for such study. Using the Treatment Evaluation Inventory (TEI) (Kazdin, 1980a), a variety of subjects have rated the acceptability of various treatments and comparisons have been made between treatments based on these evaluations. Additionally Kazdin's work has

explored the effects of several features of treatment likely to effect evaluations: severity of problems (Kazdin, 1980a), specific features of a single mode of treatment (1980b), degree of effectiveness of treatment and severity of side effects (1981a), who rates the treatment (Kazdin, French, & Sherick, 1981), and subjects and effectiveness (Kazdin, in press).

The first assumption underlying comparisons within this context is that individuals will rate each treatment presented based on the stimulus materials and react according to his or her own perceptions about that treatment. The second major assumption in this research is that the experimental design can control for such extraneous variables as the effect of rating any particular treatment before or after any other treatment. The research reported by Kazdin and his colleagues support these assumptions in a general way, and suggest that with this approach individuals can provide valid treatment acceptability ratings.

Witt and his colleagues have developed the Intervention Rating Profile (IRP) (Witt et al., 1984) and the Children's Intervention Rating Profile (CIRP) (Elliot, Witt, & Galvin, 1983, cited in, Witt & Elliot, in press) which are quite similar to the TEI developed by Kazdin, but which use language and items directed more for use with teachers and children in classroom settings. While Kazdin's studies have employed a repeated measures replicated Latin Square design, where each subject rates each treatment, Witt's studies have

employed large numbers of subjects, each rating a single case-treatment combination in a simple factorial design. In this way Witt and his colleagues have made comparisons among alternative interventions and investigated several features thought to affect treatment acceptability: potential risks to target child, teacher time required, effect on non-target children, level of skill required for implementation, materials required (Witt, Elliot, & Martens, in press; Witt et al., 1984); how treatments are presented (Witt, Moe, Gutkin, & Andrews, in press); problem type and severity (Witt, 1983); and ratings provided by children (Elliot et al., in press).

The assumptions underlying the approach employed by Witt and his colleagues are that individuals have some absolute feelings about a particular treatment combination described, and that these feelings can be quantified for the purpose or comparison using the IRP.

Several of the reported studies on treatment acceptability are suggestive of a relationship between ratings obtained and several other factors such as, teacher's level of knowledge of behavioral principles and type of case to which treatment is applied. These factors have not been satisfactorily investigated in any report to date. The implications for such investigations, particularly of the first of these factors, for teacher preservice and inservice training and for consultation are clear.

1. KNOWLEDGE OF BEHAVIORAL PRINCIPLES

The work of Jeger and McClure (1979); Norton et al. (1983); Musgrove and Harmes (1975); Frey (1970); Robinson and Swanton (1980); and McMahon, Forehand, and Griest (1981) is all suggestive of more positive attitudes toward behavioral techniques with increases in knowledge of such approaches. Several additional studies have reported empirical support for such a relationship. Wheldall and Congreve (1980a) found in a study of 116 mature inservice student teachers that the more they knew about behavior modification the more positive were their attitudes. This was also supported by a retest study following instruction in behavior modification. These findings support those of Musgrove (1974) and Throll and Ryan (1976) where positive attitudes toward behavior modification were related to higher scores on measures of knowledge about behavior modification. No significant relationship was found in any of these studies between test scores and age, experience or sex, although previous exposure to behavior modification had a highly significant effect. Hickey (1977) reports a significant and positive correlation between scores on an attitude to behavior modification scale and scores on a knowledge measure using a sample of 30 public high school counselors. Hickey found that in addition to the knowledge factor, sex and perceived socioeconomic level of school in which subjects work affected attitude ratings. Females, and teachers from higher socioeconomic schools expressed more

positive attitudes toward behavior modification. Young and Patterson (1981), in a study involving 475 undergraduate students and 67 faculty members, indicate several findings of interest. In general these findings are similar to those above with more positive attitudes toward behavior modification being associated with higher knowledge scores on a performance based measure of knowledge. However subjects who labeled themselves as informed or very informed about behavior modification did not differ from the self-labeled poorly informed and very poorly informed in their overall performance on the knowledge task. Merrett and Wheldall (1982) report that in a study involving 110 student teachers in a course in behavior modification in teaching, those students who had taken a course in behavior modification received higher grades on various measures of teaching performance and that there was a significant change in attitudes toward behavioral teaching methods.

Several other authors have alluded to such positive relationships of attitude and knowledge of or familiarity with techniques of behavior modification (Camplese, O'Bruba, & Hale, 1979; Knapp & Delprato, 1980; Lambert, 1976; Luiselli, 1981; Miller, 1981; Reppucci, 1977).

Two empirical studies report no relationship between knowledge level and attitude. Burkhart et al., (1976) in a study involving nine juvenile probation officers failed to find differences in attitude toward behavior modification on a semantic differential between a trained and control group.

The trained group scored higher on a test of behavioral competency but did not differ significantly on a [30 item, multiple choice] behavior knowledge measure. Marholin, Taylor, and Warren (1978) report on a study of 47 undergraduate students in special education which indicates no difference between high and low behavior modification test scorers on ratings of behavior modification. Furthermore, and despite the revealed differences among subjects in their knowledge of concepts, subjects did not differ in their perceptions of their competence in the use of behavior modification techniques with children.

Three issues become evident in reading this literature relating knowledge and attitudes toward behavior modification. The first of these is that while the majority of studies indicate a positive relationship, few report using validated and comprehensive measures of both knowledge and attitude. This would indicate a possible methodological shortcoming of these studies. The second issue is one expressed by Kazdin and Cole (1981) and detailed earlier: that studies of global evaluations of behavior modification fall short of ideal in that such a unitary concept may not be an appropriate target for study. The third issue is the lack of agreement among studies on the relationship which may exist. With these issues being unresolved the question of knowledge of behavioral principles being a factor in the assessment of acceptability of behavioral approaches to treatment has not been adequately researched. It appears

necessary then to investigate directly the effect of knowledge of behavioral principles as a factor in the rated acceptability of specific treatments. Such an investigation must employ a validated methodology for establishing both levels of knowledge and differences in degree of acceptability of treatments.

2. TYPE OF CASE

In 1954 the U. S. Supreme Court arrived at the decision in Brown v. Board of Education of Topeka which established the concept that educational opportunities for children are tied to success in life. In 1971 and 1972 two further landmark cases one, involving the Pennsylvania Association for Retarded Children (PARC) and the other Mills v. District of Columbia, established the rights for all handicapped children to a free appropriate public education. These decisions culminated, in 1975, in Public Law 94-142 which mandated, in detail, rights and services for all children regardless of handicap in U. S. schools (Pollack & Sulzer-Azaroff, 1981). The recent adoption in Canada of a new Charter of Rights guarantees children similar rights, although rights to services are not detailed or mandated to the same extent as in the U. S. legislation. With the advent of these legal obligations the issue of mainstreaming and providing for children the 'least restrictive' educational opportunities has developed. As a result regular classroom teachers are now challenged to meet the needs of not only a

greater number of children but also of a group of children with a greater variety of special needs. Both the general public and the teaching profession have registered particular concern about student behaviors which are disruptive and detrimental to classroom functioning and student learning (Baer, Goodall, & Brown, 1983; Lindsey & Frith, 1983). Baer et al. (1983) reviewed work by Wickman (1938) and Stouffer and Owen (1955) which suggested that teachers were mostly concerned about disruptions in classroom order (Wickman, 1938) or infractions of rules and routines and similar forms of classroom misbehavior (Stouffer & Owen, 1955). Teachers in the Baer et al. (1983) study listed physically dangerous behaviors, disruptions to learning and challenges to the teacher's authority as the three most serious problems teachers face today.

It is reasonable to assume that those students whose behaviors contribute to teachers' concerns will be the source of differential teacher reactions. Lewin, Nelson, & Tollefson (1983) and others (Brophy & Good, 1974; Silberman, 1969; 1971) have in fact presented data to suggest that classroom teachers reject those children who exhibit disruptive behaviors. Teachers, parents, and the pupils themselves don't often agree either on the seriousness of the problems or on the assignment of responsibility for the problems. Adelman, Taylor, Fuller, and Nelson (1979) found that students view their problems as less severe than their parents view the problem and that teachers rate the

students' problems even more severely than do the parents. Guttman (1982) found that pupils tend to blame all others (teachers, parents, other children, the child's environment) more than the child himself. Teachers tend to blame the misbehaving child first (psychological problems, need to let off steam, desire to gain status with classmates), and the parents second (bad example at home, parents' level of education). They tend to play down the importance of reasons associated with other children and, most pronouncedly, reasons associated with teachers themselves. Parents tend to assign blame evenly to the child himself, teachers, other children and to themselves as parents.

Since the early work of Haring (1974) suggested that teacher attitudes were instrumental in determining the future adjustment of students, many investigators have attempted to identify the sources of attitudes toward the handicapped (Brophy & Good, 1974; Foster & Ysseldyk, 1976; Hughes, Kauffman, & Wallace, 1973; Necco, 1970; Safran, Safran, & Orlansky, 1982). Investigators in this area are not in unanimous agreement regarding the effect of teacher expectancies in evaluating students' behavior. Safran et al., (1982) state that some evidence suggests that negative stereotypes bias teacher expectancies against mainstreamed students, while others disagree. Safran et al., (1982) suggest from their work that the way information about a child is presented will influence teachers' attitudes toward that child. Thus the kind of information

with which the teacher is presented and attitude of the receiving teacher will have an effect on the kinds of goals established, and possibly the methods employed in achieving those goals.

Much has been written stressing the necessity of selection of appropriate goals and the design and implementation of those goals in the most effective, efficient and professionally appropriate manner possible (Carrerra & Adams, 1970; Hochbaum, 1980; O'Leary & O'Leary, 1977; Sapon-Shevin, 1982; Sheldon-Wildgen & Risley, 1982; Stolz, 1978; Wray & Steer, 1980). The basic issue in this literature is that clients, and in particular children, have the right to the best possible assessment of needs (from appropriate perspectives) and the judicious selection of treatment procedures for the best interests of the client. Implicit in this is the understanding that individuals charged with the duty of selecting treatment strategies will in fact pay heed to differences between presenting problems and thus select treatment approaches differentially.

In the recent literature related to treatment acceptability the issues of problem type and problem severity as well as factors related to the presenting problem are not satisfactorily resolved. Lambert (1976); Elliot, Witt, & Martens (in press); Elliot, Witt, & Galvin (in press); and Wollersheim, McFall, Hamilton, Hickey, & Bordewick (1980) have found no significant effect on ratings of acceptability of treatment where different presenting

problems are considered. Gargiulo and Yonker (1983) report mixed results, however, in assessing teachers' attitudes toward the handicapped versus the non-handicapped. They found no difference between self-report measures of inservice regular, preservice regular, preservice special, and inservice special educators. However, physiological measures suggested great differences between groups, favoring more positive attitudes toward handicapped students on the part of inservice special educators.

Reflecting this contradiction are the results reported by Kazdin (1980a, 1980b, 1981a) which consistently indicate that severity of presenting problem affects ratings of acceptability of treatments. Witt, Elliot, and Martens (in press); Witt et al., (1984) found a minimal but significant effect for problem severity on acceptability of treatments. Young and Patterson (1981) found that students regarded behavior modification more appropriate for serious than for minor problems. Turkat et al., (1979) and Marholin et al., (1978) found that differential acceptance of behavior modification was attributable to problem type. Gutkin, Singer, and Brown (1980) found a moderate positive correlation between perceived problem severity and teacher preference for referral rather than consultation services. In an implementation and evaluation study of nondirective and token school programs, Mannarino and Durlak (1980) report that the type of problem presented facilitated implementation of services.

The work of Kazdin and his colleagues (Kazdin, 1980a, 1980b, 1981a, in press; Kazdin et al., 1981) in acceptability of treatments has in all cases employed two stimulus case descriptions in order to "assess the possibility that treatment evaluations of the students were based upon, or restricted to, unique characteristics of the stimulus material" (Kazdin, 1980a, p. 263). In this regard Kazdin refers to the arguments presented by Maher (1978) which caution simply that "if we wish to generalize to populations of stimuli, we must sample from them" (p. 646). In an attempt to sample different populations of stimuli Kazdin has employed cases which vary on the important dimensions of psychiatric or behavioral problems; settings where problems occur; as well as age, gender and intellectual characteristics of the cases described. Based on the Maher (1978) work criticism might be leveled at the approach Kazdin has employed in that the critical stimulus variables may not be sampled and secondly that such "scripts" (Maher, 1978, p. 695) may in fact not represent 'real' people. In other words, these may not be a sample from the population to which generalization is to be made. Maher suggests that our hesitation in generalizing from a single stimulus case is increased substantially by the prospect of generalizing from a case that is not known to have existed at all.

In a study assessing the acceptability of time-out procedures Norton et al., (1983) substantially replicated

earlier work by Kazdin (1980b). Although in contrast to the Kazdin study, Norton et al. (1983) controlled for particular case stimuli and found age of child described affected ratings of both effectiveness and acceptability. In support of Kazdin's findings, however, other stimulus features such as sex of the child, and location where behavior occurred did not affect ratings of acceptability or effectiveness.

It is still unclear, however, whether such stimulus features as written background information will affect teachers ratings of the acceptability of interventions. Safran et al., (1982) using several case types found that background information on hearing impaired and normal children affected teachers' perceptions of student behaviors. However, this finding did not extend to the case of the acting-out child. It is necessary, then, to investigate further the effect of background information on teachers' ratings of the acceptability of treatments, in particular to establish these effects in light of the present emphasis on mainstreaming of children who exhibit a variety of problem behaviors.

3. METHOD OF ASSESSING ACCEPTABILITY

The two predominant reported methods for rating the acceptability of treatments are the Treatment Evaluation Inventory (TEI) (Kazdin, 1980a) and the Intervention Rating Profile (IRP) (Witt, Elliot, & Martens, in press). A third method described in a recent study (Norton et al., 1983)

used a single item rating scale to assess the relative acceptability of various time-out procedures. While this study employed a relatively simple measure, the results support those of Kazdin's (1980b) study. These findings would suggest that investigations of the methodology for assessing acceptability may be fruitful.

In many classrooms and clinical consultation or treatment settings the individual charged with selecting a particular treatment for implementation does so not by considering the acceptability of a single treatment approach in isolation, but rather by comparing and contrasting a number of possible options and selecting that one which appears to have the 'best fit', given the present situation. The assumption here is that an individual confronted with a problem situation (which is multi-dimensional) will employ some implicit personal criteria to evaluate all alternatives and select an appropriate treatment. The range of alternatives which would be considered would firstly be restricted by the individual's knowledge of alternatives and secondly by the criteria used in evaluating those known alternatives.

While it may be reasonable to accept the above assumption, both the Witt and the Kazdin studies have gone to considerable lengths to avoid the very process of direct comparisons among treatment alternatives. Witt has his subjects look at a single treatment only and makes comparisons among alternative treatments using grouped data.

Kazdin has each subject evaluate each of three or four treatments but attempts to provide statistical control through experimental design for any carryover (or comparison effect). Both of these approaches have yielded valuable information toward our understanding of several dimensions of treatment acceptability but it is important to consider whether these or some other approach might generate more relevant and generalizable data. An alternative approach might allow for the evaluation of treatment alternatives through the direct and explicit comparison of treatments along similar dimensions as investigated in the previous research. Such an approach has not received attention in the research literature to date and an appropriate methodology has not as yet been developed.

D. PROBLEM SUMMARY

In the review of literature related to treatment acceptability three problem areas were identified. There exists a lack of systematic investigation of the effect of knowledge of behavioral principles on teachers' ratings of the acceptability of selected treatment alternatives, although research suggests a relationship. Present research is also inconclusive regarding the effect of type of case on such ratings. While research evidence does exist substantiating the comparison of acceptability of various treatment alternatives it is not clear whether the present methodology is either the most efficient or is intended to

be a close approximation of the clinical decision-making process.

E. PURPOSE OF THE STUDY

Reflecting the problems identified, the purpose of the present study was to question the effect on the acceptability of alternative classroom treatment procedures of certain experimental and subject variables. The initial question addressed in the study was to determine whether practicing elementary school teachers differentiate, in terms of acceptability, alternative classroom treatment procedures which might be applied with children's problem behaviors. The second question was whether level of knowledge of behavioral principles affects the way in which individuals assess the acceptability of selected classroom treatment procedures. The third question was whether certain stimulus characteristics of the presented case description affect the acceptability of treatment ratings which teachers provide. The fourth question was whether an alternative method of assessing acceptability will reflect the findings established using the TEI. It was the final intent of this study to establish data from a sample of local teacher subjects which might be compared, in a general way, with data established in similar investigations carried out in other settings.

F. SIGNIFICANCE OF THE STUDY

Classroom teachers appear to have been given or to have taken a major responsibility for identifying children who are exhibiting problem behaviors, of initiating assessment procedures and selecting and implementing treatments appropriate to the setting and to the child so identified. The quality of implementation and even the selection of appropriate procedures appears to depend a great deal upon the individual teacher's knowledge of and views toward the array of possible alternatives.

Beyond the fact that teachers' perceptions of classroom interventions are valuable in their own right, the relation between these perceptions and other variables is also of some importance (Garfield, 1983). An individual's subjective evaluation of a treatment may effect whether it is implemented properly, whether it is effective (or perceived to be effective), the length of time it will be used or whether it will even be used at all. Witt & Elliot, in press, p. 25)

Although the study of treatment acceptability has received some recent research attention, major research questions and problems remain. One need is for data on the generality of findings on differential evaluations of treatment procedures on the basis of acceptability. If, using the present research methodology, it can be established that individuals (in this case teachers) do differentiate procedures, then it may be possible to identify salient variables related to the procedures, or perhaps to the individuals evaluating treatments. Particular treatments or features of some treatments may then be identified and amended in ways which might enhance their acceptability without compromising their

effectiveness. It may be also that identification of certain subject variables may suggest changes in future directions for curriculum and practicum experiences for student teachers, as well as inservice training for practicing teachers.

The majority of published research to date on treatment acceptability has employed undergraduate and graduate psychology students as subjects. It must be established, therefore, whether present or potential consumers of classroom treatment procedures (e.g., teachers) respond in a fashion similar to that established in other studies by other populations.

In this study an attempt was made to analyse consumers' ratings of acceptability of treatments and to identify other features of both treatments and subjects which affect those ratings. By examining other factors such as individuals' degree of knowledge of behavioral principles, or level of training, the study may suggest ways in which teacher preparation, treatment presentation, or in fact, treatment design might be altered. The goal of such alterations would be the development and use of effective and acceptable treatments.

Ideally, in the study of treatment acceptability, the body of knowledge ought to be established by having real teachers who are facing real problems with real children select from among all the alternative treatment procedures those procedures which they feel 'fit' best. All such

selections could be recorded including any salient problem, teacher, or subject variables. The application of these procedures could be monitored and studies carried out on such effects as integrity of application, effectiveness of procedure, and effect on non-target children, to name a few. In this way it might be established over a long period of time and probably at great expense, more direct and believable evidence both about acceptability of treatments and the many factors which contribute to treatment acceptability. Although this approach may be ideal in terms of the wealth of data it would produce, the research costs would seem to be prohibitive. The next best approach then must be that approach which most nearly approximates the ideal (i.e., the most naturalistic) but which does not have the attendant prohibitive costs. It may be that a more naturalistic procedure for assessing acceptability of treatments can be devised which will yield useful data at no greater research costs than are presently expended in the study of acceptability of treatments.

It is important and worthwhile, particularly in any new research area, to establish a body of data on subject and setting generality. Thus systematic replication (Sidman, 1960) of current research is a viable and important purpose of study. This study will attempt to establish through an examination of the questions raised above whether findings developed in diverse other settings and using other types of subjects (Kazdin, 1980a, 1980b, 1981, in press; Kazdin et

al., 1981; Witt & Martens, in press; Witt, Moe, Gutkin, & Andrews, in press) can be generalized to other populations and settings.

II. METHODOLOGY

The purpose of this chapter is to describe the method followed to address the research questions posed in the previous chapter. The chapter begins with a broad overview of the nature of the study, followed by more complete descriptions of the instrumentation and experimental design, the subjects of the study, and the experimental and data preparation procedures. Finally, an outline of the data analysis procedure is presented. The results of these analyses follow in Chapter Three.

A. NATURE OF THE STUDY

This is primarily a causal-comparative study but makes use of procedures of experimental and analytic survey methodologies. Briefly, a screening sample of volunteer elementary school teachers completed an instrument intended to assess their level of knowledge of behavioral principles as applied to children. A study sample of regular class elementary school teachers was selected from the screening sample on the basis of the knowledge measure results and the demographic data supplied. Study subjects were presented with one of two child case descriptions which differed in terms of the history of treatment the child had experienced. One child was described as having a history of special class assignment with intensive teacher support in a structured setting. In the description of the other case no mention is made of prior treatment. Subjects were presented with

written descriptions of four treatments which might be appropriate for application in a classroom with a child such as the one presented in the case description. Following the reading of each treatment, subjects rated the acceptability of the treatment on a four part treatment acceptability measure. When all four treatment ratings were completed subjects responded to a final measure designed to solicit reactions to the case description, the relative acceptability of the treatments and subjective responses to treatment needs for the child described.

The primary focus of the study was upon differences in treatment ratings provided by teachers, with respect to the teachers' knowledge level, the case described and the particular treatments presented.

B. INSTRUMENTATION

In addition to a demographic questionnaire completed by all subjects, the present study employed three measurement instruments: an initial knowledge measure, a four part measure of treatment acceptability, and a final questionnaire intended to solicit other information relevant to the case description and to treatment ratings. A description of each of the instruments and the method of scoring subject responses follows.

1. DEMOGRAPHIC QUESTIONNAIRE

The demographic questionnaire (see Appendix A) was completed by all teachers prior to all other study procedures. Characteristics surveyed included subjects' sex, age, level of educational degree or training, years of teaching experience, and grade level of children taught. Subjects were also asked to indicate whether they had any special training in child management, and to further categorize any such experiences. Teachers' responses on the questionnaire were intended to provide a means of both characterizing the study population and screening subjects on the basis of required characteristics (i.e., regular classroom teachers status).

Scoring the Demographic Questionnaire

Teachers' responses to the questionnaire were coded for sex, years of experience, age in years, grade level taught, and level of professional education (teacher certification only, earned degree plus certification, graduate degree plus certification). Responses to the final question of special training were coded as yes or no.

2. KNOWLEDGE MEASURE

A modified (16-item) version of the Knowledge of Behavioral Principles as Applied to Children (KBPAC) (O'Dell, Tarler-Benlolo, & Flynn, 1979) was used to assess understanding and application of basic behavioral principles with children. In the development of the KBPAC, 70 items

were generated based on behavioral principles found expressed in four widely recognized texts on the management of children's behavior (Becker, 1971; Hall, 1971; Patterson & Gullion, 1968, Patterson, 1971). These 70 items were administered to a sample of 147 persons representing both lay and professional people with a wide variety of experience with behavior modification. O'Dell et al. report a Kuder-Richardson reliability coefficient of 0.94 and an odd-even split-half correlation of 0.93 for this version of the KBPAC. Fifty items which attained the highest point-biserial correlation with total score (all > 0.30 , mean=0.49) were retained for the final version. While items in the final version did not cover all principles found in the cited texts the authors suggested that the 50 items together provided a reasonable sampling of the content of interest.

O'Dell et al. report using the KBPAC in two studies involving volunteer subjects in a 5-hour child management training workshop. In both studies odd-even split-halves of the instrument were administered pre and post training. The first group of volunteer parents demonstrated an increase from 48% to 85% correct on the alternate form. Four other samples of undergraduate university subjects, one third of whom were psychology students, were provided with the same training experience. This additional sample of 91 respondents increased pre/post training from 57% to 85% correct using the alternate form.

A modified 45-item version of the KBPAC was used in a study in which the effect of incorporating formal training in social learning principles in behavioral parent training was investigated (McMahon, Forehand, & Griest, 1981). Five items from the KBPAC which did not reflect content appropriate to the purpose of the study were not used in the modified version. In this study the KBPAC reflected significant differences in the desired direction between groups who had training only and training plus instruction in social learning principles. Other studies have reported development of instruments similar in purpose and design to that of the KBPAC (Marholin, Taylor, & Warren, 1978; Milne, 1982). The finding of similar results in these studies suggests the validity of the approach used in the KBPAC to assess knowledge of behavioral principles as applied to children.

Furtkamp, Giffort, and Schiers (1982) have examined two 10-item parallel forms of the KBPAC for use in research or evaluation settings. These shortened versions were developed in reaction to the length of time (30-60 minutes) required for most subjects to complete the full KBPAC. A sample of 164 Mental Health Technician trainees completed the 50-item KBPAC. From the results of this sample item point-biserial correlations were used to determine the 20 items which best predicted total score for the test. Overall, the mean item point-biserial correlation was 0.296 (S.D.=0.157) and the median was 0.297. The range 0.693 (from -0.153 to =0.540).

The twenty items selected ranged from 0.540 to 0.383. These items were then randomly assigned to form the two ten-item versions. These versions were administered to a sample of 111 trainees following a course in basic behavior modification, and item and test analysis completed on their responses. The authors found that reducing the length of the test had little impact on the internal consistency of the tests (Kuder-Richardson 20 of 0.862 for the 50-item version, and 0.735 and 0.765 for the two ten-item versions). The correlation between the two forms, determined from the responses of Mental Health Technician trainees, was 0.65 ($r^2=0.423$, $F(1,109)=80.17$, $p < 0.001$). When corrected for attenuation, the estimated correlation of true scores on the tests is 0.87. Furtkamp et al. (1982) concluded that research results supported use of these two versions as parallel forms.

For the purpose of the present study a further revised version of the KBPAC was developed. Based on the finding of O'Dell et al. (1979) and Furtkamp et al. (1982), a 16-item version of the KBPAC was prepared. While Furtkamp et al. (1982) report satisfactory results from the two 10-item versions, they also suggest that some of the items used in the two forms may need reworking. Although all items discriminated in the right direction several attained point-biserial correlations less than 0.30 with the total test. Additionally, both 10-item versions attained reliabilities below that for the 50-item version (0.86) and

below the desired level of 0.80. Application of the Spearman-Brown formula (Nunnally, 1970) with the reliabilities obtained in the Furtkamp et al. study (1982) indicated that if test items homogenous with either version were added to increase test length to 15 or 16 items, the desired level of reliability might be attained (for 16 items Form A reliability=0.82, for 16 items Form B reliability=0.84). From the 20 items retained by Furtkamp et al. (1982), those 16 items which exhibited point-biserial correlation coefficients greater than 0.30 with total test score were selected for the new version of the KBPAC. This selection eliminated three items from Form A and one from Form B (point-biserial correlations of 0.19, 0.22, 0.27 and 0.29).

The newly revised form together with the two 10-item versions developed by Furtkamp et al. (1982) were administered to samples of students (n=95) registered in five undergraduate educational psychology classes at the University of British Columbia. Within class, each version of the revised KBPAC was randomly distributed so that approximately equal numbers of subjects completed each version. Each subject completed one version only. Administration time for the 10-item versions was approximately 4-5 minutes, the 16-item version taking 3-4 minutes longer.

Student responses to each version of the KBPAC were analysed using the Laboratory of Education Research Test

Analysis Package (LERTAP) (Nelson, 1974). This test analysis procedure provides both item and total test analysis. Examination of results from this analysis revealed that all item point-biserial correlations for the 16-item KBPAC exceeded 0.30 (range 0.31 to 0.69, mean .51). Item results of the 10-item versions were less satisfactory (Form A range 0.23 to 0.63, mean .42, and Form B, range 0.06 to 0.72, mean .46). These latter results were consistent with the findings of Furtkamp et al., with the same items achieving low point-biserial correlations with total test. Results of the LERTAP analysis for the full test are presented in Table 1. Results of Furtkamp et al. (1982) are included for the purpose of comparison.

The results summarized in Table 1 indicate an internal consistency reliability for the 16-item version which surpasses that of either of the 10-item versions in the pilot study. The obtained standard error of measurement for the two 10-item versions is not substantially different from that reported by Furtkamp et al. (1982) although the variability among subjects is much less in the pilot sample. While it might be expected that a longer version of the test would demonstrate an improved reliability, the results obtained for the two 10-item versions in the pilot study are barely adequate (.42 and .58) in comparison to that obtained for the 16-item version. Additionally, the estimated reliabilities for each of the 10-item versions inflated by the Spearman-Brown formula (10a, $r=.54$; 10b, $r=.69$ when

TABLE 1
KBPAC PILOT STUDY RESULTS

Version	Number of Subjects	Mean	Standard Deviation	Reliability	Standard Error of Measurement
<u>Pilot</u> ¹					
10A	31	5.87	1.91	0.42	1.31
10B	31	5.10	2.15	0.58	1.32
16	33	8.18	3.90	0.81	1.63
<u>Furtkamp et al.</u> ²					
10A	111	4.78	2.62	0.74	1.35
10B	111	4.78	2.75	0.77	1.33
25 odd	164	12.14	4.61	0.77	2.20
25 even	164	13.17	4.17	0.74	2.15
50	164	25.31	8.26	0.86	3.07

Note. Reliabilities reported for pilot study are Hoyt's, those for Furtkamp et al. (1982) are Kuder-Richardson. For dichotomously scored items such as those represented here, these two forms are equivalent.

¹Pilot subjects were student teachers.

²Study subjects were Mental Health Technician trainees.

increased in length to 16 items) still do not reach the desired level of reliability, and certainly do not improve upon the result obtained using the 16-item version. In comparing the obtained results with those reported by Furtkamp et al. (1982) the 16-item version most nearly approximates the desired level of internal consistency reliability. While the reliability obtained for this form (.81) with a sample of student teachers is less than that

reported for the 50-item version (.86) with the Mental Health Technician trainees, use of the 16-item version represented a potential saving of 20 to 40 minutes of subject time. Since it was a desire of the author to obtain an accurate measure of subjects' knowledge of behavioral principles as applied to children in the most efficient manner, the 16-item version of the KBPAC was selected as the appropriate measure for inclusion in the study.

A copy of the 16-item version of the KBPAC, including instructions for completion, is presented in Appendix B.

Scoring the Knowledge Measure

The KBPAC responses for both the instrument development and screening samples were hand scored and total score for each subject calculated. Item responses received a value of one for agreement with keyed response and zero for any other response. In addition, subject item responses were coded and placed on a computer file with 100% verification. No errors were found on verification. This computer file was used for verification of total score as well as for completion of item and test analysis using the LERTAP computer program.

3. TREATMENT ACCEPTABILITY MEASURES

The treatment acceptability measures consisted of the Treatment Evaluation Inventory (TEI) (Kazdin, 1980a) and a three part (Evaluative, Potency, Activity) Semantic Differential (SD) (Osgood, Suci, & Tannenbaum, 1957).

a. Treatment Evaluation Inventory

The TEI consists of 15 Likert-type items which together reflect an individual's overall evaluation of the acceptability of a particular treatment. Included are such considerations as the treatment's overall acceptability; whether it should be recommended for broad application; how fair, humane or effective it might be; whether it would be appropriate for use with someone not able to give consent and whether it fits with common notions of what treatment should be.

Regarding development of the TEI, Kazdin (1980a) reports that 45 original items were generated which appeared to relate to client evaluation of treatment. From this number Kazdin reports that 16 were selected which best related to treatment of children and the use of punishment. Respondents were asked to rate a particular described treatment on each of the 16 items using a seven point Likert-type scale. One item, for example, asked respondents to indicate the acceptability of treatment on the dimension of 'suitability' of the treatment for application with problems other than those described. The anchor points, not suitable at all to very suitable, for example, represent ratings of treatment on one aspect of acceptability. Subsequent factor analytic studies conducted to assess the evaluative validity of the TEI led to deletion of one of the 16 items. The 15-item version of the TEI used in the present study (see Appendix C) was developed and implemented in

several studies subsequent to Kazdin's first (1980a) study.

Scoring the TEI

All 15 items on the TEI are responded to by subjects checking one position on a seven point scale which most nearly represents the subject's response to the described treatment. For the purpose of scoring, each response position was given a numerical value, where a value of one was assigned to the negative anchor point (e.g., not at all suitable) and a value of seven assigned to the positive anchor point (e.g., very suitable) for each of the 15 questions. A value of four represented a neutral position on a particular item. Thus a total acceptability score could range from a minimum of 15 to a maximum score of 105.

Individual subject responses to each item for each treatment were coded and input to a computer file with 100% verification. Fewer than 5% errors in coding and input were found; all were corrected prior to analysis. The LERTAP computer program was used to calculate individual acceptability scores for each treatment rated, and to provide item and test analysis data.

b. Semantic Differential

The three part SD used in the study (see Appendix D) consisted of 15 items, including five from each of the evaluative, potency and activity dimensions described by Osgood et al. (1957). The version used in the study is one developed by Kazdin (1980a) who gave two reasons for

including the SD with the TEI as measures of treatment acceptability. The first was to increase the number and diversity of variables for the subsequent factor analysis, allowing for a more careful delineation of a homogeneous factor. The second was to provide a partial validation of the TEI. It was expected that the evaluative dimension would correlate more highly than the potency and activity dimensions with the TEI. Both the TEI and the evaluative dimension of the SD were thought to reflect overall acceptability while the potency and activity dimensions were thought to reflect important dimensions of treatment perhaps not directly relevant to the evaluation of acceptability.

Scoring the SD

The 15 items on the SD are responded to by subjects placing a check mark in one of seven positions relative to each set of bipolar adjectives. For the evaluative dimension a value of seven was assigned to the check space nearest to the positive pole (e.g., good) and a value of one to the negative pole (e.g., bad). Positions between were assigned values accordingly. For the five items representing the potency dimension, the pole representing greater potency (e.g., strong) was assigned a value of seven and the opposite pole (e.g., weak) a value of one. For the activity dimension the more active pole (e.g., fast) was assigned a value of seven, the less active (e.g., slow) a value of one. Scores for each dimension were considered separately so that for each subject three dimension totals were found for

each of the treatments rated. Total scores for each dimension could range from a minimum of five to a maximum of 35 for each treatment.

As with the TEI, item responses for each individual under each treatment condition were coded and input to a computer file for calculation of total score, item, and test analysis statistics using the LERTAP computer program. Subject responses were 100% verified, with fewer than 5% errors in coding and input to the data file. All errors were corrected prior to analysis.

c. Psychometric Properties

Kazdin (1980a) piloted the treatment acceptability measures (the 16-item TEI and the 15 items of the SD) with 60 undergraduate psychology students. In the pilot study subjects each heard one of four treatments descriptions, each presented as the treatment might be applied to a clinical case previously described to the subject. Subjects then rated the acceptability of the treatment using the measures described above.

The scores for pilot subjects were submitted to a principal components factor analysis which was rotated to the varimax criterion (Kazdin, 1980a). The results of this analysis revealed two factors. Fifteen of the sixteen items of the TEI loaded highly on the first factor (pattern coefficients ranged from .61 to .95), as did the evaluative items from the SD. Loading for the potency and activity dimension of the SD were low (less than .40) for the single

factor that characterized items of the TEI. These data confirmed the expectation that the TEI assessed overall evaluations of the treatments and suggested that these evaluations were made independent of both potency and activity ratings. Additional support of this analysis is reported in a second administration to 144 college students enrolled in undergraduate psychology courses (Kazdin, 1980a). In this study items from the TEI yielded high factor loadings on the first factor (pattern coefficients ranged from .56 to .95) and items from the evaluative dimension of the SD loaded highly (pattern coefficients ranged from .69 to .89) on this factor as well.

Treatment acceptability measures used in subsequent studies (Kazdin, 1980b, 1981a, in press; Kazdin et al., 1981) included the three part SD and a TEI consisting of the 15 items which loaded on the first factor of the analyses reported above. Results of analysis of data from these latter studies supported the original findings. Although no reliability data on either the TEI or the dimensions of the SD were reported in these studies, a high internal consistency may be inferred from the results of the factor analysis (i.e., items loaded together on a common factor) and the range inter-item correlations for items of the first factor of the TEI (from .35 to .96, median $r = .67$).

Validity of the TEI for assessing treatment acceptability is partly established above and further supported by subsequent studies in which the TEI was used as

a dependent measure (Hobbs, Walle, & Caldwell, 1984; Kazdin, 1980b, 1981a, in press; Kazdin et al., 1981). Results of these studies established the ability of the TEI to distinguish, in the expected direction, between several separate treatment strategies across a variety of populations and settings.

4. CASE AND TREATMENT QUESTIONNAIRE

The final measure used in the study was a questionnaire developed for the purpose of the present study. This questionnaire (see Appendix E) was intended to solicit information from teachers about certain features of the case and of the treatments which had been rated. In this regard four questions were posed.

The first question asked subjects to rate the perceived seriousness of the case described. Kazdin (1980a) found that treatments were generally rated more positively (i.e., were assigned higher acceptability scores) when treatments were presented for application with more severe cases than those for less severe cases. The first question was intended to provide an indication of the degree of perceived severity of the case presented and to detect potential differences in severity of cases described.

Question two required subjects to indicate from their own experience how typical the case described was of students in the regular classroom. This question was intended to obtain some indication of incidence and possibly

to lend validation to the plausibility of teachers having to confront this kind of problem. Since treatment descriptions relate directly to the case description, it was important that teachers see the case as a 'typical' difficult child or at least a plausible one so that treatment ratings might better reflect acceptability in a relevant and plausible context. As Maher (1978) suggests, one danger in using scripts (here the case description) is that, in an attempt to represent an ideal 'typical' case, a "case is created that like the manticore, may never have existed in nature" (p. 645). Teacher responses to this question indicate whether in fact the assumed case description is relevant and realistic enough to allow teachers to relate to treatment descriptions.

The third question required subjects to consider all four treatments at one time and to assign each to relative ranks on the basis of how likely the subject would be to use each in the classroom. If the TEI procedure can be perceived to reveal some absolute measure of an individual's ratings of the acceptability of a particular treatment, then this particular question asks the subject to consider the relative acceptability of the four treatments described. Witt and Elliot (in press) stated that a "reason for assessing acceptability is to increase the likelihood that a treatment will be used and will be implemented with integrity" (p. 2). They suggest the need for research to link acceptability and use, but also stress the need to

establish procedures for assessing acceptability which are relevant to the selection and use of treatments in the school (and other clinical) setting(s). It is important to develop a procedure which as nearly as possible approximates the clinical decision-making process, whereby the 'most appropriate' or 'best available' of several known options is selected for use in the particular situation. This is particularly important in light of research which suggests that some teachers find no intervention which is totally acceptable (Witt & Elliot, in press). This question is one attempt at establishing a methodology for soliciting such information.

The final question required teacher subjects to describe the approach they would most likely use if the child described were in their class. In this question teachers were not restricted by the treatments which had been presented although all had been exposed to the descriptions prior to answering this question. A major intent of this question was to provide some data for future treatment acceptability research. Teacher responses to this question may reveal approaches to classroom intervention upon which acceptability information ought to be accumulated. Witt and Elliot (in press) and Lambert (1976) suggest that teachers are typically aware of only one or two treatments. It is important then to consider the dimension of acceptability of those known treatments relative to other possible treatments. This question may reveal also the

breadth of approaches seen by teacher subjects as appropriate for use in the classroom setting.

Scoring the Case and Treatment Questionnaire

In response to the first question, teachers indicated which scale point most nearly represented their estimate of the seriousness of the problems described. The positions on the five point scale ranged from one (not very serious at all) to five (extremely serious). The second question, which asked teachers to describe how typical was this child, did not produce responses which were readily quantifiable as no metric was provided. Therefore responses were coded dichotomously. Responses which indicated that the child was somewhat typical (e.g., "one every year", "two like him in my class right now", "one out of fifty", and "I've seen four or five in 10 years of teaching") were scored as one. Responses which indicated that the case described was not at all typical were coded as zero.

Responses to question three on the questionnaire represented subjects ranking of the relative likelihood of their using each of the described treatments. Subjects indicated their ranking by arranging treatment code letters in order from 'most likely to use this one' to 'least likely to use this one'. For the purpose of scoring these responses, positions in the arrangement were assigned a value ranging from, one for 'least likely' to four for 'most likely'.

Question four was not scored for direct analysis in the study as it did not yield data that was readily coded. Moreover, these responses were not of direct interest to the purpose of the study.

Coded responses were added with 100% verification to the computer file containing the coded responses to the treatment acceptability measures. As described above for scoring of treatment acceptability measures these data were scored using the LERTAP computer program. The Case and Treatment Questionnaire data were treated as precoded subtests with zero weighting for this analysis, allowing for tabulation of responses without affecting the analysis of the treatment acceptability measures.

C. DESIGN OF THE STUDY

1. DEPENDENT VARIABLES

The dependent variables in the present study represent subject scores on the TEI and the three dimensions of the SD for each of the treatments. The treatments included: Medication (M), Time Out (TO), Reinforcement (R), and Positive Practice (PP). The dependent variables were measured as the total score attributed to each treatment on each of the treatment acceptability measures. With regard to the ranking procedure in the Case and Treatment Questionnaire, the dependent variable was the relative rank assigned to each treatment. A more detailed description of

procedures for scoring and assigning values is presented in the description of data preparation.

2. INDEPENDENT VARIABLES

This study included three independent variables which are represented as factors in the experimental design. These were knowledge group, type of case, and type of treatment. Subjects' scores on the knowledge measure were used as an initial blocking factor in the study. Details regarding scoring of the knowledge measure are included in the description of data preparation. On the basis of total score on the knowledge measure subjects were separated into two groups using the median scores as the 'cutpoint'. Subjects who scored at the median were not considered further. A more detailed description of this procedure is presented in the section on assignment of subjects.

The second independent variable, type of case, had two levels corresponding to the history of special class treatment attributed to the case (see Appendix F). Both descriptions are of the same child and are presented in the same way. Mike R. is described as a young boy who is exhibiting severe behavioral problems in school and in a variety of other settings. The description of Mike S., the second case, is the same as that of Mike R. except for the addition of an introductory paragraph which presents Mike S. as coming into the present teacher's class as a result of school and district efforts toward mainstreaming. His

history states that he has been in a special class setting where he has been used to individual attention and special programming. Other than the addition of this paragraph on the history of treatment the descriptions of the two cases are identical and are adaptations of descriptions used by Kazdin (1981a).

The third independent variable is treatment type. Four treatments, Medication, Time Out, Positive Practice, and Reinforcement (Kazdin, 1981a) were considered. The treatment descriptions, presented in Appendix G, are representative of reported versions of interventions applied to the type of target behaviors presented in the case description (Kazdin, 1981a, p. 497).

The design of the study necessitated consideration of the effect of two additional independent variables. These were the variables related to sequence in which the treatments appeared (row effect) and the order, or relative position in which a treatment appeared (column effect). These might be distinguished from the foregoing independent variables as nuisance or design variables.

3. EXPERIMENTAL DESIGN

The design used in the present study was an extension of the 4 x 4 Latin-square design (Myers, 1979) and is similar to that used by Kazdin (1981a) of which this study is, in part, a systematic replication. Since four treatments were to be presented and repeated measures taken on the

dependent variables, treatment presentation was randomized in the square following the procedure for randomization of Latin-squares described by Kirk (1968) and by Cochran and Cox (1950). One of four standard squares was selected at random. Rows and columns were then randomized independently yielding the particular square used in the present study. Within this square each of the four treatments occurs once in each order or column (i.e., first, second, third, or fourth) and once in each sequence or row. The Latin-square then represents a random selection of four of the 24 possible sequences in which the four treatments could be presented.

In addition to the treatment variable, the factors of knowledge level and type of case were included in the design as between-subjects factors, hence the Latin-square was replicated once at each level of each of these factors yielding a $2 \times 2 \times (4 \times 4)$ (knowledge level \times case level \times (sequence \times order)) replicated Latin-square design. The layout of the overall design is presented in Figure 1.

Myers (1979) has suggested that Latin-square designs have several potential advantages over other designs. First, the Latin-square design allows for the investigation of several variables with less expenditure of time and fewer subjects than a comparable factorial design. Additionally, Myers suggests that an even more important advantage is the efficiency of the Latin-square design relative to other designs. In using the Latin-square in a repeated measures

FIGURE 1
LAYOUT OF THE EXPERIMENTAL DESIGN

Knowledge Group	Case	Sequence	Order 1st	2nd	3rd	4th
-----------------	------	----------	--------------	-----	-----	-----

High

Mike S.

1	M	TO	R	PP
2	TO	PP	M	R
3	PP	R	TO	M
4	R	M	PP	TO

Mike R.

1	M	TO	R	PP
2	TO	PP	M	R
3	PP	R	TO	M
4	R	M	PP	TO

Mike S.

1	M	TO	R	PP
2	TO	PP	M	R
3	PP	R	TO	M
4	R	M	PP	TO

Low

Mike R.

1	M	TO	R	PP
2	TO	PP	M	R
3	PP	R	TO	M
4	R	M	PP	TO

design, the researcher is able to remove error variance due not only to individual difference, as in a regular repeated measures design, but also sequence and order or temporal effects. This advantage is particularly important when repeated measures on the dependant variable are likely to be affected by carryover effects of potent treatments.

D. PROCEDURE

1. SUBJECTS

The subjects of this study were 107 volunteer practicing elementary school teachers, solicited from nine schools in the Burnaby School District. Burnaby is a major suburban center within the Greater Vancouver area. It is representative along most relevant dimensions of the surrounding urban and suburban area which comprises a large proportion of the teaching and student population of the province. Since a major focus of the study was ratings provided by regular classroom teachers, the final subject sample was restricted to teachers who had a regular class enrolled and had no more than three hours each week as assigned special education or administration time.

Subjects were solicited through an oral presentation to school staff meetings by the researcher and facilitated through the cooperation of the school district research committee and the local school administration. The introductory presentation and subsequent procedures required

approximately 10-15 minutes of subject time. Teachers were presented with a written introduction and statement of purpose and procedures for the study. All staff indicating interest in participation completed the demographic questionnaire and the required consent form. Following this, teachers responded to the 16-item knowledge measure.

One hundred and seven teachers agreed to participate in the study and completed the knowledge measure and demographic questionnaire.

The mean score on the knowledge measure for the initial or screening sample of 107 teachers was 7.03 (S.D.=3.68) with a maximum score of 16 possible. This sample of 30 male and 77 female teachers had a mean age of 39 years and a mean of 14 years of teaching experience. The characteristics of the teachers in the screening sample are summarized in Table 2.

All teachers indicating interest in participation in the study had the opportunity to complete all experimental procedures. Sixty-four of those who met study requirements and indicated willingness to participate were designated within the study as subjects and had their data treated accordingly. Data provided by other participants were retained and used in analyses outside the direct purpose of the study. This included analysis of various properties of the measurement instruments used in the study.

TABLE 2
CHARACTERISTICS OF SCREENING SAMPLE

Age (years)		Experience (years)	Knowledge Level
Mean	39.29	14.16	7.03
S.D.	8.74	7.52	3.68
Range	32-60	0-35	0-16 (max. 16)

Sex	Educational Levels	Levels of Class Enrolled
Male=30	Teach. Cert.=10	Primary=42
Female=77	Deg. + Cert.=83	Intermediate=45
	Grad. Deg. + Cert.=14	Special=20

Schools Represented	Total Subjects	Special Training
n=9	n=107	Yes=56; No=51

2. ASSIGNMENT OF SUBJECTS

Those teachers who scored at the median level (mdn=7) on the knowledge measure (n=13), or who did not meet the stated criterion (i.e., regular classroom teacher) (n=12) or who did not wish to participate further in the study (n=9) were not considered for the purpose of selection of

subjects. Of those remaining, 32 subjects from each of the knowledge groups (high, low) were randomly selected for assignment to treatment conditions. Within each knowledge group four teachers were randomly selected and assigned to each of eight experimental conditions (two child case descriptions x four treatment sequences). Thus, a total of 32 teachers were randomly selected from the remaining pool of subjects at each knowledge level for participation and data collection, and identified within the study as the subject sample.

3. DATA COLLECTION

Experimental procedures for this second meeting proceeded, as the first, with group administration. However, for this session teachers received individual packages of materials appropriate to the preassigned experimental condition. Differences in packages represent the two different case descriptions and four different sequences in which treatment were presented.

Each subject was provided with an evaluation package containing the appropriate case description, four treatment descriptions, four copies of the treatment acceptability measures and a single copy of the case and treatment questionnaire. The case description included and the sequence in which treatments were arranged was determined prior to subjects viewing materials, in the manner described earlier.

After reading their assigned case description teachers were asked to consider the first treatment in their package, and then to complete the treatment acceptability measures with reference to their case. These materials were then put aside and the procedure repeated for the remaining three treatment descriptions. Subjects were asked to consider each treatment separately, but to remember that each treatment related to the particular case description presented.

Following completion of the four separate treatment ratings subjects were asked to complete the case and treatment questionnaire. The second session, including giving of instructions, treatment evaluations, and final questionnaire required approximately 30 minutes of subject time. The experimenter was present during both this and the first session, to answer questions and to ensure compliance to experimental procedures.

E. DATA PREPARATION AND ANALYSIS

1. DATA PREPARATION

Scoring procedures for subject data follow those described earlier under Instrumentation. All data were coded and entered onto computer files with 100% verification. All errors were corrected prior to analysis.

2. ANALYSIS OF DEMOGRAPHIC QUESTIONNAIRE

The responses of teachers to the demographic questionnaire provided data to describe the final sample of 64 teachers. Descriptive statistics are provided for the final sample of 64 teachers, separated into the two knowledge groups. The SPSS:X computer program CROSSTABS (SPSS Inc., 1983) was used to provide summary descriptive statistics of these two groups.

3. PSYCHOMETRIC ANALYSIS

a. Knowledge Measure

The psychometric analysis of subject responses to the KBPAC was accomplished using the LERTAP computer program. This program provides subjects scores, mean, standard deviation, estimate of internal consistency reliability (Hoyt, 1941), and standard error of measurement, as well as statistics for item difficulties and point-biserial correlation coefficients.

Following satisfactory results of the psychometric analysis and for the purpose of identifying high and low scoring subject pools, the median of the distribution of subjects scores was calculated. This score then provided the 'cut point' for dividing the two groups, with those subjects scoring at the median eliminated from assignment to subject pools.

b. Treatment Acceptability Measures

The first analysis of subject responses on each of the dependent measures was an item and test level analysis. Once again the LERTAP program was used to obtain means, standard deviations, standard errors of measurement, as well as item-test correlations and Hoyt estimates of the internal consistency reliability of the tests. In the present study ratings for each treatment on each of the dependent measures were treated as separate subtests in the LERTAP analysis. The appropriateness of summing item responses within each dependent measure was determined by an inspection of item and test statistics provided in the analysis. The total score obtained for each subject on each of the dependent variables was used in the second analysis of treatment ratings.

4. ANALYSIS OF TREATMENT ACCEPTABILITY RATINGS

The analysis of treatment acceptability ratings was for significant effects of the independent variables of knowledge level, child case, and treatment type, based on teacher responses on the dependent measures. Additionally, the analysis considered the effects of several other variables which are functions of the experimental design (subjects, sequence, and order). All of these effects were investigated by utilizing a replicated Latin-square analysis of variance with repeated measures.

Although Latin-square designs have certain benefits and advantages which were described earlier (c.f., experimental design) such designs also have attendant costs in terms of the amount of information available from the analysis. The major sacrifice in the analysis of such designs stems from the fact that information about interactions between treatment variables and row and/or column variables are impossible to extract. It is possible, for example, for treatments to be affected differentially by either row or column placement or a combination of both. Such interactions, if they do exist may obscure the main effects of interest.

Within this type of design treatments are said to be aliased within specific higher order interaction effects that are themselves assumed to be zero (Kirk, 1968; Rosenthal & Rosnow, 1984; Winer, 1962). In the present analysis treatment effects are aliased within the sequence x order (row x column) interaction term. While sequence, order, and sequence x order interaction effects may be of some interest in themselves, the primary focus of the research is the aliased treatment effect. It is possible, in experimental designs such as the present one, to extract treatment effects (and interaction effects of treatment with independent variables other than sequence and column) from the specific higher order interactions. These interactions may, however, contain in addition to the treatment effect, variance attributable to sources other than treatment. In

the present analysis the sequence x order interaction term (which is the treatment alias) may contain treatment effect plus a 'Latin Square Residual' (LSR) (G. J. Johnson, personal communication, August, 1984). The LSR may be both isolated and tested for significance, but where significant will in most cases be uninterpretable. Myers (1979) suggests that a significant LSR can be interpreted to be an effect of a treatment x order interaction (or a treatment x order x other independent variable interaction, depending on the LSR under consideration) if there is strong a priori evidence for assuming that no sequence interactions are present. Myers further states, however, that, "there is no test that permits determining which interaction component is present [when the LSR is significant, and concludes that]. . . such a result only reveals that some interaction component is present" (p. 282).

In several studies using a similar experimental design (Kazdin, 1980a, 1980b, 1981a, in press; Kazdin et al., 1981), the exact model of analysis is not stated and complete source tables for the analysis of variance are not presented. Kazdin (in press) suggests a two stage analysis; the first of these being a 3 x 3 (sequence x order) Latin-square analysis of variance for repeated measures. Kazdin states that this stage of analysis is used to evaluate the impact of the way in which treatments were presented. The second stage of analysis focuses on the remaining variables of interest, ignoring sequence and order

factors.

In earlier work, Kazdin (1980a) appears to follow the approach suggested by Myers (1979) of testing effects of the LSR. In the first study by Kazdin using this design (1980a), tests of treatment effects on the potency and activity dimensions of the SD were made against the usual within subjects error terms, but the tests of treatment effects for the TEI and the evaluative dimension of the SD were made against the LSR as an error term, leading to a negatively biased and hence highly conservative F test (Kazdin, 1980a, p. 265; see also Myers, 1979, p. 276). Kazdin reported using that approach because the LSR obtained was larger than the within subjects error term. Kazdin does not report testing the LSR for significance in this analysis, but a significant result may be inferred. In experiment two of the Kazdin (1980a) study the initial analysis for sequence and order effects also required use of the conservative F tests (larger mean square and fewer degrees of freedom in design) indicating the likely presence of a significant LSR in this analysis as well. Other studies using the same basic experimental design, reported by Kazdin and his colleagues (Kazdin, 1980a, 1980b, 1981a, in press; Kazdin et al., 1981), however do not report results of analysis requiring such tests.

A further methodological issue is raised in reviewing the two stage analysis reported in the Kazdin studies. While Kazdin states, for example, that there is "no direct

interest in evaluating characteristics of the case and its interactions with other variables" (1981a, p. 496), it is apparent that neither Case Description nor Treatment Efficacy (the other between-subjects variables in this study) were considered in the first stage of analysis. Analysis at this first level then does not account for any variance due to difference between levels of these factors or interactions involving either or both. In the second stage of analysis the effects of the between-subjects variables and treatments are investigated but the effects of sequence and order are not.

The experimental design used in both the present study and those reported by Kazdin and his colleagues has the potential for identifying several sources of variance not investigated in the two stage analysis. The analysis selected for the present study was a single stage analysis which accounts for all identifiable sources of variance in the experimental design. Computational formulae used in the analysis are attributable to Johnson (G. J. Johnson, personal communication, August, 1984) and are exactly equivalent to Myers (1979).

Employing the approach described by Myers (1979, p. 279-285) for investigating additional independent variables by 'Latin squaring treatment combinations', the computer program BMDP:8V (Dixon, 1983) was used to provide four separate analyses of data corresponding to the ratings from each of the four treatment acceptability measures. The

analysis approach was the same for each and proceeded as follows. Since the design of the present study required analysis of a Latin-square with repeated measures, two analyses of variance were performed and then combined to yield the sources of variance corresponding to the analysis proposed by Johnson (G. J. Johnson, personal communication, August, 1984) and Myers (1979).

The first repeated measures analysis of variance was performed using subject data presented in column order. The second repeated measures analysis of variance was performed with the same data reordered in treatment level ordering. The second analysis (treatment level ordering) provided the sums of squares for treatment, treatment x knowledge interaction, treatment x case interaction, and treatment x knowledge x case interaction necessary to decompose all levels of the sequence x order alias. These results were then combined with the results of the first analysis to produce the ANOVA table described by Johnson (G. J. Johnson, personal communication, August, 1984) and Myers (1979).

In this analysis the LSR is tested for significance. Given a significant LSR the tests of relevant main effects are made against the LSR rather than the usual (within subjects) error term. Given significant F ratios in these analyses, the means were compared using Tukey's Honestly Significant Difference (HSD) test (Bruning & Kintz, 1977).

5. ANALYSIS OF CASE AND TREATMENT QUESTIONNAIRE RESPONSES

Only one the four questions posed in the questionnaire yielded results which could be analysed. This analysis was for the effect of differences in ranking of treatments provided in question three. This analysis utilized both Friedman's Two-Way Analysis of Variance by Ranks and Kendall's Coefficient of Concordance. This analysis was performed by the computer program BMDP:3S (Dixon, 1983). In this analysis ranks given to each treatment were entered as the dependent variable. Given a significant Friedman test statistic, Nemenyi's (1963) procedure was used to compare means.

F. CHAPTER SUMMARY

The primary purpose of this chapter was to describe the method followed in the study. The method was first described in general terms as the nature of the study, then more specifically. The development, selection, and implementation of four measurement instruments was described. The presentation of the design of the study included both a description of the variables of interest and the particular experimental design selected. The procedure section described the subjects of the study and the procedures whereby subjects were assigned to conditions and data gathered. The final section of the chapter presented the methods of data preparation and analysis. The methods

followed in both the analysis of the psychometric properties of the measurement instruments and analyses of the demographic, and case and treatment questionnaires were presented. The development and selection of appropriate procedures for the analysis of the treatment acceptability ratings were described in some detail.

III. RESULTS

The intent of this chapter is to present the results of the analysis as described in Chapter Two. The chapter begins with a description of the subjects who contributed responses in the study, followed by the results of the psychometric analysis. The next and major portion of the chapter is devoted to the presentation of results of the analysis of treatment acceptability ratings. The results of the analysis of responses to the case and treatment questionnaire are presented last. The chapter concludes with a summary of the results focused on results related to the primary hypotheses.

A. SUBJECTS

Subjects for the study were volunteer regular class elementary school teachers who were selected and assigned within the study following the procedures outlined in Chapter Two.

Briefly, 32 teachers were selected for each of the (high and low) knowledge groups within the study. The study sample of 64 subjects did not appear to differ from the screening sample on the characteristics reported on the demographic questionnaire. Table 3 presents a summary of the characteristics of the groups of high and low knowledge teachers (as measured by the KBPAC). Significance tests between the two knowledge groups for each of the reported characteristics indicated no significant differences between

TABLE 3
SUBJECT DEMOGRAPHIC CHARACTERISTICS

Characteristic	Screening Sample	Subject Sample	Low Knowledge	High Knowledge
Subjects Total	107	64	32	32
Sex				
Male	30(28%)	12(19%)	9(28%)	3(9%)
Female	77(72%)	52(81%)	23(72%)	29(91%)
Age				
Mean	39.29	39.17	39.81	38.53
S.D.	8.74	8.82	9.59	8.09
Experience (years)				
Mean	14.16	13.73	15.06	11.78
S.D.	7.52	6.77	8.40	4.51
Educational Level				
Teach. Cert.	10(9%)	7(11%)	5(16%)	2(6%)
Deg. + Cert.	83(78%)	54(84%)	27(84%)	27(84%)
Grad. Deg. + Cert.	14(13%)	3(5%)	0	3(9%)
Knowledge Level				
Mean	7.03	7.00	3.84	10.16*
S.D.	3.68	3.80	1.92	2.27
Child Management Training				
Yes	56(52%)	33(52%)	12(38%)	21(66%)**
No	51(48%)	31(48%)	20(63%)	11(34%)

* $\phi = .82$, $p < .05$

** $\chi^2 = 4.00$, $p < .05$

knowledge groups for sex, age, number of years of teaching experience, and level of educational training. As would be expected, the two knowledge groups do differ significantly ($\phi = .82$, $p < .05$) with regard to score on the knowledge measure. Additionally, significant differences were found

between knowledge groups with regard to training in child management ($\chi^2 = 4.00$, $p < .05$), with the high knowledge group having a greater proportion of subjects with training than not, and the low knowledge group having a greater proportion of subjects without training than with. Included in Table 3 as well are summaries of the characteristics of both the screening and total subject samples, affording comparison and contrast between groups along each of the reported dimensions.

B. PSYCHOMETRIC ANALYSIS

1. KNOWLEDGE MEASURE

The summary of results from the item analysis of the KBPAC administered to the screening sample are presented in Table 4. Examination of the individual items revealed that the difficulty indices ranged from .16 to .70, and the item-test point-biserial correlations ranged from .35 to .62. The corresponding Hoyt estimate of reliability, reported in Table 5, was .78. Results of the total test analysis for the screening sample are summarized in Table 5.

2. TREATMENT ACCEPTABILITY MEASURES

a. Treatment Evaluation Inventory

Since each subject used the TEI to respond to each of the four treatments and responses to each were considered to represent separate tests, item-test correlations were

TABLE 4
KBPAC ITEM ANALYSIS

Item	Difficulty Index	Point Biserial Correlation with Total Test Score
1	.28	.39
2	.51	.35
3	.16	.55
4	.22	.49
5	.54	.45
6	.60	.58
7	.70	.41
8	.33	.62
9	.50	.57
10	.24	.52
11	.64	.56
12	.43	.43
13	.44	.43
14	.48	.40
15	.38	.57
16	.58	.50

calculated separately for each treatment. Examination of the point-biserial correlations revealed that for the most part, items were positively related to test score.

Fifty of the items exceeded .50, of the remaining only one was negative (-.07). These results compare favorably with those presented and discussed in Chapter Two.

The results of the test analysis for each application of the TEI appear in Table 6. As shown by the value of the Hoyt's estimate of internal consistency, reliability of measurement was high.

TABLE 5
KBPAC TEST ANALYSIS

Number of Individuals = 107	Number of Items = 16.00
Mean = 7.03	Highest Score = 16.00
S.D. = 3.68	Lowest Score = 0.00

Hoyt Estimate of Reliability = 0.78

Standard Error of Measurement = 1.66

b. Semantic Differential

Item and test level analyses were performed on responses to each of the subtests representing the three dimensions of the SD for each treatment. Results of this analysis indicate that of the 60 items considered only one did not attain a point-biserial correlation greater than .20. This was item number 5 on the activity dimension within the Reinforcement condition, representing the adjective pair ferocious/peaceful. In this case the subtest correlation attained was .178, lower than the other three conditions (Medication=.66, Time Out=.55, and Positive Practice=.47). Since values attained in the other conditions were acceptable the result in the Reinforcement condition was considered insufficient evidence to reject this item. Results of the item level analysis in general supported use of all items in calculation of total score for each

TABLE 6

TEI TEST ANALYSIS

Treatment	Mean	S.D.	High Score†	Low Score‡	Hoyt Estimate of Reliability	Standard Error of Measurement
M	50.97	18.01	93	18	0.94	4.35
TO	69.97	20.00	96	16	0.95	4.24
R	79.03	14.22	103	52	0.89	4.64
PP	66.45	19.48	103	18	0.94	4.68

n=64

Note. M = Medication; TO = Time Out; R = Reinforcement;

PP = Postive Practice

†Maximum highest score = 105

‡Minimum lowest score = 7

dimension of the SD.

The results of the test level analyses provided through the LERTAP program are summarized in Table 7. In all cases the data represent scores from 64 subjects on five questions for each subtest. The maximum rating therefore would be 35 and a minimum rating would be five. As summarized in Table 7, Hoyt's estimates of reliability were high for each dimension of the SD over each treatment condition. The results reported in Tables 6 and 7 supported the use of the measurments employed. Given these findings, subjects' total

TABLE 7
SEMANTIC DIFFERENTIAL TEST ANALYSIS

Treatment	Mean	S.D.	Hoyt Estimate of Reliability	Standard Error of Measurement
<u>EVALUATIVE DIMENSION</u>				
M	17.50	7.46	0.96	1.38
TO	23.48	6.83	0.90	1.97
R	27.05	6.05	0.91	1.59
PP	21.94	6.61	0.91	1.80
<u>POTENCY DIMENSION</u>				
M	23.38	6.79	0.93	1.62
TO	23.03	5.68	0.90	1.64
R	19.25	5.39	0.86	1.79
PP	20.63	6.43	0.95	1.32
<u>ACTIVITY DIMENSION</u>				
M	19.95	7.44	0.88	2.34
TO	21.83	6.48	0.88	1.99
R	20.42	4.49	0.75	1.99
PP	22.30	5.83	0.85	2.03

Note. M=Medication; TO=Time Out; R=Reinforcement;
PP=Positive Practice

score on each of the dependent measures for each treatment were considered in subsequent analyses.

C. ANALYSIS OF TREATMENT ACCEPTABILITY RATINGS

1. TREATMENT EVALUATION INVENTORY

The results of the analysis of variance of the TEI ratings are presented in Table 8. Tables 9, 10, and 11 present comparisons between means for the a posteriori analysis, following significant F tests of overall differences reported in Table 8.

As shown in Table 8 there were five significant sources of variance in this analysis. Two of these, LSR_{CFB} and LSR_{CDB} correspond to differential carryover or environmental effects on treatment ratings and are not interpreted directly. Given significant Latin square residuals, however, corresponding interaction effects (i.e., AF, AD) were tested against the LSR rather than the usual within subjects error term (C subjects within FDB) (Myers, 1979). Discussion of the remaining sources, knowledge, treatment and order follow, beginning with the knowledge factor.

The significant F ratio for the knowledge group factor [$F(1,48)=6.33$, $p<.02$] indicated that knowledge group membership had a significant effect on ratings of treatment acceptability on the TEI. In general, subjects in the high knowledge group gave higher ratings ($\bar{M}=69.38$) across treatments than did the low knowledge group ($\bar{M}=63.88$). The factor of knowledge group is not present in any significant interaction effects (with the exception of the residual interaction term LSR_{CFB}).

TABLE 8
ANOVA FOR TEI

Source	df	MS	F
Knowledge (F)	1	1897.74	6.33*
Case (D)	1	3.33	0.01
Sequence (B)	3	376.16	1.26
FD	1	73.27	0.24
FB	3	321.26	1.07
DB	3	75.79	0.25
FDB	3	330.10	1.10
Subjects within FDB	48	299.81	
Order (C)	3	1101.35	3.68*
CF	3	51.50	0.17
CD	3	131.51	0.44
Treatments (A)	3	8751.84	29.27*
CB Latin Square Residual (LSR_{CB})	6	299.83	1.00
CFD	3	193.66	0.65
AF	3	574.43	0.82
LSR_{CFB}	6	697.84	2.33*
AD	3	422.00	0.64
LSR_{CDB}	6	658.31	2.20*
AFD	3	185.58	0.62
LSR_{CFDB}	6	268.88	0.90
C Subjects within FDB	144	298.98	

* $p < .05$

The significant F ratio for treatment [$F(3,144)=29.27$, $p<.001$] indicates a significant difference between levels of the factor attributed to differences in treatment ratings. The treatment factor does not participate in any significant interaction effects tested. Results of comparisons between treatment rating means on the TEI using the Tukey Honestly Significant Difference (HSD) procedure are summarized in Table 9. The summary in Table 9 shows that the Reinforcement intervention was rated as significantly more acceptable than the other interventions using the TEI. Although the Positive Practice and Time Out procedures were not rated significantly different from each other on the TEI both were less acceptable than Reinforcement and more acceptable than the Medication intervention. The Medication intervention was rated significantly less acceptable than all other treatment interventions.

The significant F ratio for order of presentation indicates that TEI treatment ratings differed as a result of the order in which treatments were rated. This result must be interpreted with some caution in light of the two significant residual terms (LSR_{CFB} and LSR_{CDB}) involving the order factor. Table 10 presents a summary of comparisons between order means on the TEI ratings. The results presented in Table 10 indicate that, in general, treatment ratings were not significantly different in first and second positions. However, ratings in second position were higher than those in third and fourth position. Similarly ratings

TABLE 9

TUKEY (HSD) COMPARISONS BETWEEN TREATMENT MEANS ON THE TEI

Treatment	Mean	R	Difference		M
			TO	PP	
Reinforcement (R)	79.03	--	9.06*	12.58*	28.06*
Time Out (TO)	69.97		--	3.52	19.00*
Positive Practice (PP)	66.45			--	15.48*
Medication (M)	50.97				--

* $p < .05$

in the first position were, in general, higher than those in the fourth position. It must be noted that perhaps these are general trends and that the presence of significant LSR effects suggests that individual knowledge groups (in the case of the term LSR_{CBF}) and individual case groups (LSR_{CBD}) did not necessarily follow this specific trend. It may be that the presence of the LSR effects suggests a significant treatment x order interaction with knowledge group, and with case (Myers, 1979). There is no test to substantiate directly such an interpretation. However, LSR effects can be displayed in order to aid in the interpretation of the main effect for the order factor. Two graphs are presented to illustrate the complexity of the order effect and to reinforce the caution in its interpretation. Figure 2 presents LSR effects for the Reinforcement and Positive

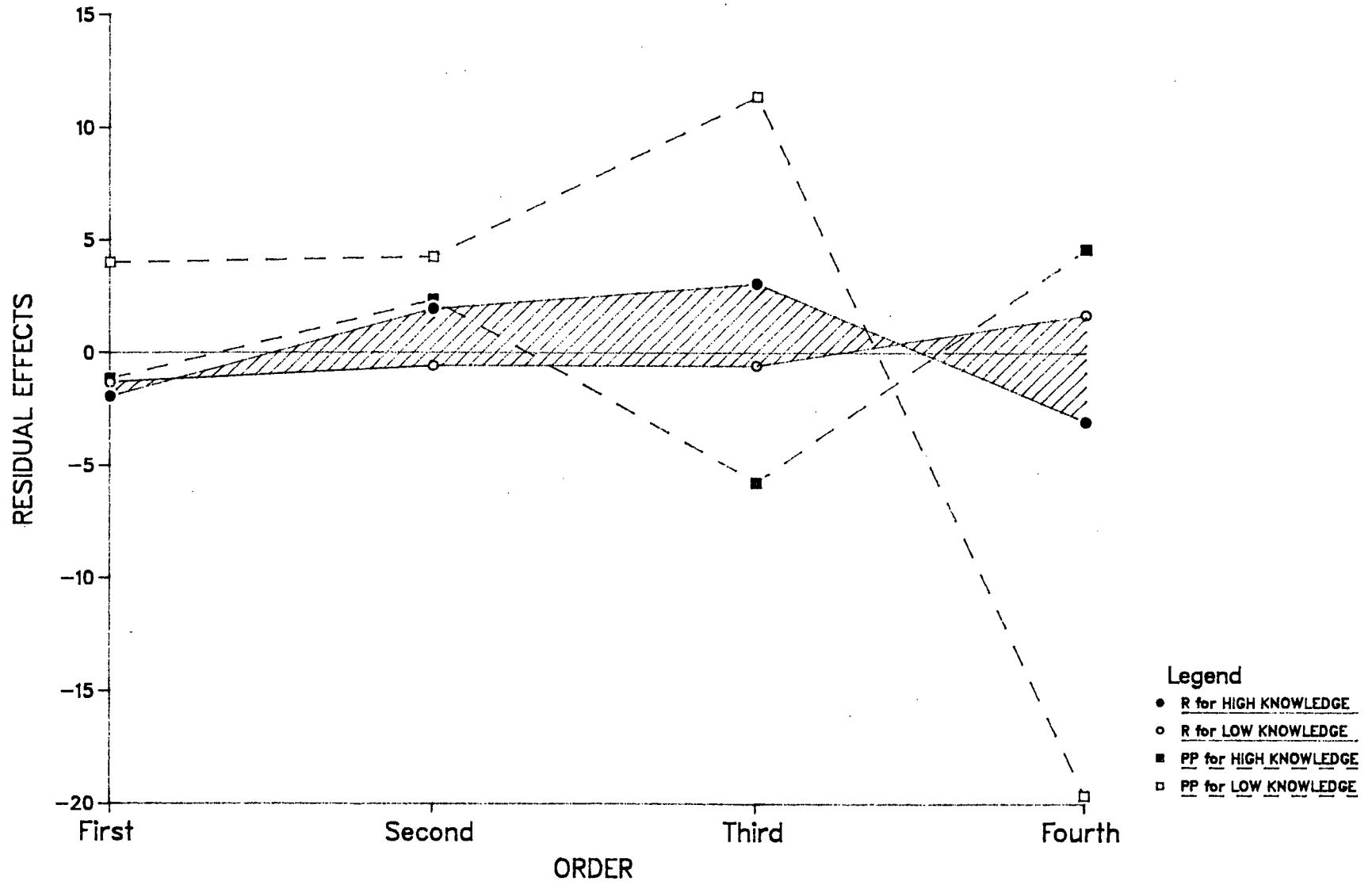
TABLE 10
TUKEY (HSD) COMPARISONS BETWEEN ORDER MEANS ON THE TEI

Order	Mean	Difference			
		2	1	3	4
Second (2)	72.05	--	4.72	7.41*	9.64*
First (1)	67.33		--	2.69	4.92*
Third (3)	64.64			--	2.23
Fourth (4)	62.41				--

* $p < .05$

Practice interventions at each of the four orders for the two knowledge groups. As can be seen in Figure 2, the ratings for the Reinforcement intervention are fairly stable across orders for the low knowledge group, and only slightly less so for the high knowledge group. The shaded area highlights the residual effects for the two knowledge groups on the Reinforcement intervention. In stark contrast the two broken lines represent the LSR effects for the Positive Practice intervention, and appear to reflect a different pattern of effects. The major difference between the two treatments is that the Positive Practice intervention ratings appear to be affected by the order factor in a different way than the Reinforcement intervention ratings. Secondly, with respect to the Positive Practice intervention ratings, the two knowledge groups appear to be affected

Figure 2
LSR EFFECTS—CxFxB RESIDUALS
Reinforcement and Positive Practice
for High and Low Knowledge



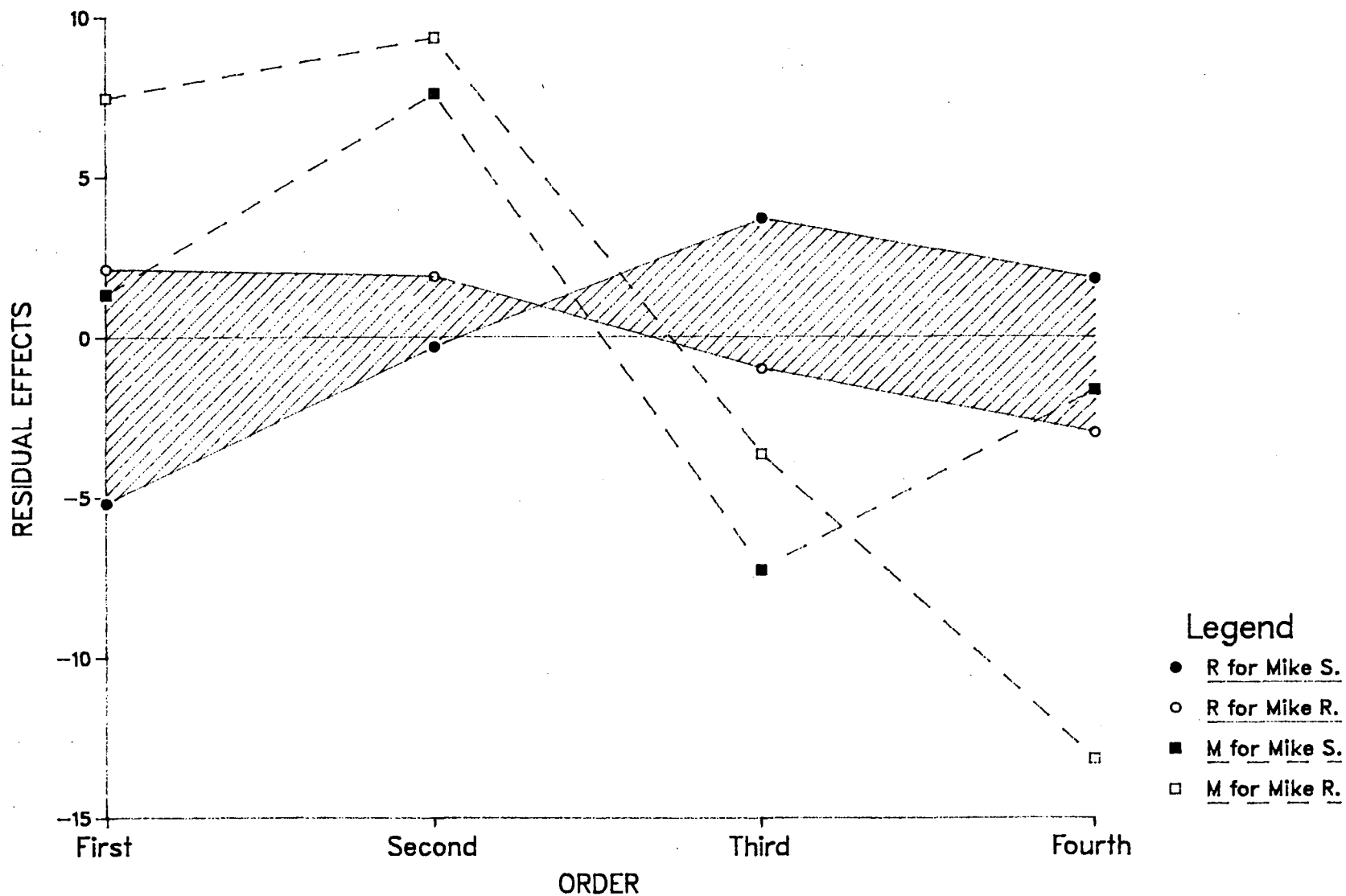
quite differently by the order factor.

A similar contrast can be drawn in a comparison of the LSR effects of two treatments for each of the two cases. Figure 3 presents the LSR effects for the Medication and Reinforcement interventions for the cases of Mike S. and Mike R. Here again the Reinforcement intervention appears to be more stable across orders for subjects assigned to the case of Mike R. than for those assigned to Mike S. As well the two case groups appear to be affected somewhat differently by the order factor. For the Mike R. group ratings are higher in the first than in fourth position. For the Mike S. group the opposite is true. The ratings of the two case groups on the Medication intervention appear to be affected similarly in first, second, and third order but differ in degree when in fourth. Additionally the Medication intervention ratings appear to be affected by the order factor in a different way than are the Reinforcement intervention ratings.

Since the order x treatment interaction in each of the LSR effects is perfectly confounded with the sequence x treatment interaction further interpretation of the effect of order of presentation and of the residual effects is unwarranted.

Thus the significant, and interpretable, effects in the analysis of TEI acceptability ratings were for knowledge level and treatment type. The high knowledge group teachers rated treatments generally more positively than did those

Figure 3
LSR EFFECTS—CxDxB RESIDUALS
Reinforcement and Medication
for Two Cases



teachers in the low knowledge group. This finding was not affected by the case which the teachers considered. The second major finding was that teachers did differentiate treatments in terms of acceptability ratings on the TEI. In this context, the Medication intervention was rated lowest in overall acceptability, while the Reinforcement intervention received the highest ratings of the four treatments presented. The Positive Practice and Time Out interventions did not differ in their acceptability, but both were rated higher than Medication although lower than Reinforcement.

2. SEMANTIC DIFFERENTIAL

As described in Chapter Two, teachers' responses for each dimension of the SD were considered separately in the analysis. The results of analyses for each dimension of the SD are presented in the same way as the TEI results.

a. SD Evaluative Dimension

The results of the analysis of variance for the Evaluative dimension are presented in Table 11. As shown in Table 11 there were three significant sources of variation in treatment acceptability ratings on the Evaluative dimension of the SD: treatment, sequence and order. The result of primary interest to the study is the significant effect for treatment; therefore, this result and comparisons of treatment means are presented first. The presentation of significant sequence and order effects follows.

TABLE 11
ANOVA FOR SD EVALUATIVE DIMENSION

Source	df	MS	F
Knowledge (F)	1	87.89	2.10
Case (D)	1	4.00	0.10
Sequence (B)	3	124.28	2.98*
FD	1	5.06	0.12
FB	3	67.42	1.61
DB	3	18.03	0.43
FDB	3	5.49	0.13
Subjects within FDB	48	41.77	
Order (C)	3	161.32	3.70*
CF	3	23.40	0.54
CD	3	2.89	0.07
Treatments (A)	3	1001.80	22.95*
CB Latin Square Residual (LSR_{CB})	6	39.22	0.90
CFD	3	9.45	0.22
AF	3	62.90	1.44
LSR_{CFB}	6	62.87	1.44
AD	3	19.51	0.45
LSR_{CDB}	6	92.12	2.11
AFD	3	27.01	0.62
LSR_{CFDB}	6	65.64	1.50
C Subjects within FDB	144	43.64	

* $p < .05$

The significant F ratio for treatment [$F(3,144)=22.95$, $p<.001$] indicates a significant difference between ratings given particular treatments. The absence of interactions involving treatments suggests that the differences in ratings were consistent across levels of knowledge group factor as well as with different cases. Results of the post hoc comparisons between treatment means using Tukey's HSD procedure are summarized in Table 12. The results summarized in Table 12 show that the Reinforcement intervention was given ratings significantly higher than the other three interventions. The ratings given the Time Out and the Positive Practice interventions did not differ significantly; however, both were rated significantly higher than the Medication intervention. The Medication intervention was rated significantly lower than all other treatment interventions.

Tables 13 and 14 present comparisons of sequence and order means respectively. The results of the comparisons between means for the order of presentation summarized in Table 13 indicate that ratings given in the third position were generally lower than those given to treatments in the second or first position but not significantly different from those given in the final position. Ratings given in the second position were also higher than those in the fourth position. The summary of results of comparisons between sequence means in Table 14 reveals that the treatments in sequence four were rated higher than those in either

TABLE 12

TUKEY (HSD) COMPARISONS BETWEEN TREATMENT MEANS ON THE SD
EVALUATIVE DIMENSION

Treatment	Mean	Difference			
		R	TO	PP	M
Reinforcement (R)	27.05	--	3.57*	5.11*	9.55*
Time Out (TO)	23.48		--	1.54	5.99*
Positive Practice (PP)	21.94			--	4.44*
Medication (M)	17.50				--

* $p < .05$

sequence two or three. No other differences were found between sequence means.

The presence or order and sequence effects in this analysis while not of primary interest, suggests that carryover effects may be present in the analysis. The major concern of such a result is that these effects may obscure some other effects of interest (i.e., knowledge, case, and treatment differences).

To summarize the results of the analysis of the SD Evaluative dimension, the single interpretable effect of interest was that of treatment type. This finding reflected the result of the TEI analysis. Subjects in the two different knowledge groups and reading different case descriptions did not rate treatments differently.

TABLE 13
TUKEY (HSD) COMPARISONS BETWEEN ORDER MEANS
ON THE SD EVALUATIVE DIMENSION

Order	Mean	Difference			
		2	1	4	3
Second (2)	24.52	--	1.52	3.24*	3.35*
First (1)	23.00		--	1.72	1.83*
Fourth (4)	21.28			--	0.11
Third (3)	21.17				--

* $p < .05$

TABLE 14
TUKEY (HSD) COMPARISONS BETWEEN SEQUENCE MEANS
ON THE SD EVALUATIVE DIMENSION

Sequence	Mean	Difference			
		4	1	2	3
Four (4)	24.30	--	1.58	2.35*	3.30*
One (1)	22.72		--	0.77	1.72
Two (2)	21.95			--	0.95
Three (3)	21.00				--

* $p < .05$

b. SD Potency Dimension

The results of the analysis of variance for the potency dimension are presented in Table 15. The results summarized in Table 15 show treatment to be the only significant source of variation in treatment ratings on the potency dimension of the SD. A summary of comparisons between treatment means for this measure is presented in Table 16. As shown, the Reinforcement intervention received significantly lower potency ratings than any of the other interventions. Medication, which received the highest potency rating was significantly higher than either the Reinforcement or Positive Practice interventions but not significantly different from the Time Out intervention. The Time Out intervention was rated higher on the potency dimension than both the Positive Practice and Reinforcement interventions.

c. SD Activity Dimension

The results of the analyses of variance for the activity dimension of the SD are presented in Table 17. The results of the analyses of variance for the activity dimension of the SD show that order of presentation was the single significant effect. Comparisons between ordered presentation means on the activity dimension are summarized in Table 18. The results of comparisons between means for order of presentation on the activity dimension show that treatments when presented in the third position are given higher activity ratings than in any of the other positions. No other differences were found.

TABLE 15
ANOVA FOR SD POTENCY DIMENSION

Source	df	MS	F
Knowledge (F)	1	13.14	0.28
Case (D)	1	45.56	0.98
Sequence (B)	3	109.09	2.34
FD	1	9.00	0.19
FB	3	3.55	0.08
DB	3	56.03	1.20
FDB	3	36.01	0.77
Subjects within FDB	48	46.66	
Order (C)	3	26.48	0.74
CF	3	45.92	1.28
CD	3	6.76	0.19
Treatments (A)	3	248.93	6.93*
CB Latin Square Residual (LSR_{CB})	6	8.56	0.24
CFD	3	55.80	1.55
AF	3	51.14	1.42
LSR_{CFB}	6	32.28	0.90
AD	3	44.60	1.24
LSR_{CDB}	6	47.69	1.33
AFD	3	0.75	0.02
LSR_{CFDB}	6	7.70	0.21
C Subjects within FDB	144	35.94	

* $p < .05$

TABLE 16
TUKEY (HSD) COMPARISONS BETWEEN TREATMENT MEANS
ON THE SD POTENCY DIMENSION

Treatment	Mean	Difference			
		M	TO	PP	R
Medication (M)	23.38	--	0.35	2.75*	4.13*
Time Out (TO)	23.03		--	2.40*	3.78*
Positive Practice (PP)	20.63			--	1.38*
Reinforcement (R)	19.25				--

* $p < .05$

3. SUMMARY OF TREATMENT ACCEPTABILITY RESULTS

The results of the analyses of variance for the treatment acceptability measure revealed several significant findings.

The effect of knowledge group was significant for acceptability ratings on the TEI [$F(1,48)=6.33$, $p < .05$] only. The effect of treatments was, however, statistically significant for acceptability ratings on the TEI [$F(3,144)=29.27$, $p < .05$] and for the Evaluative [$F(3,144)=22.95$, $p < .01$] and Potency [$F(3,144)=6.93$, $p < .05$] dimensions of the Semantic Differential. The results of the comparison tests among treatment means are summarized in Table 19.

TABLE 17
ANOVA FOR SD ACTIVITY DIMENSION

Source	df	MS	F
Knowledge (F)	1	6.89	0.16
Case (D)	1	50.77	1.14
Sequence (B)	3	59.84	1.35
FD	1	1.56	0.04
FB	3	38.32	0.86
DB	3	46.59	1.05
FDB	3	13.72	0.31
Subjects within FDB	48	44.42	
Order (C)	3	102.45	2.83*
CF	3	33.63	0.93
CD	3	0.32	0.01
Treatments (A)	3	79.69	2.20
CB Latin Square Residual (LSR_{CB})	6	29.30	0.81
CFD	3	13.53	0.37
AF	3	20.64	0.57
LSR_{CFB}	6	50.60	1.40
AD	3	11.14	0.31
LSR_{CDB}	6	37.23	1.03
AFD	3	69.13	1.91
LSR_{CFDB}	6	32.59	0.90
C Subjects within FDB	144	36.26	

* $p < .05$

TABLE 18
TUKEY (HSD) COMPARISONS BETWEEN ORDER MEANS
ON THE SD ACTIVITY DIMENSION

Order	Mean	Difference			
		3	1	4	2
Third (3)	22.91	--	1.99*	2.16*	2.99*
First (1)	20.92		--	0.17	1.00
Fourth (4)	20.75			--	0.83
Second (2)	19.92				--

* $p < .05$

On both the TEI and SD Evaluative, the Reinforcement intervention received the highest acceptability ratings, while the lowest acceptable ratings were assigned to the Medication intervention. In terms of potency these results are reversed. Order was significant for the TEI [$F(3,144)=3.68$, $p < .05$] and for the Evaluative [$F(3,144)=3.70$, $p < .05$] and Activity [$F(3,144)=2.83$, $p < .05$] dimensions of the Semantic Differential. These results are summarized in Table 20.

The order effect was not consistent in favoring one position over the three measures. Additionally, lack of clear differences across dependent measures does not allow any conclusion regarding this effect.

TABLE 19
SUMMARY OF TUKEY (HSD) COMPARISONS BETWEEN TREATMENT
CONDITIONS

Dependent Measure	Condition †			
Treatment Evaluation Inventory	R	<u>TO</u>	<u>PP</u>	M
SD Evaluative Dimension	R	<u>TO</u>	<u>PP</u>	M
SD Potency Dimension	<u>M</u>	<u>TO</u>	PP	R

Note. Any two means underlined by the same line are not significantly different, whereas any two means not underlined by the same line are significantly different. All differences are at the $p < .05$ level. Means are arranged in descending order from left to right.

† R = Reinforcement; PP = Positive Practice; TO = Time Out; M = Medication.

TABLE 20
SUMMARY OF TUKEY (HSD) COMPARISONS AMONG ORDERS OF
PRESENTATION

Dependent Measure	Orders			
Treatment Evaluation Inventory	<u>Fourth</u>	<u>Third</u>	First	Second
SD Evaluative Dimension	<u>Third</u>	<u>Fourth</u>	<u>First</u>	<u>Second</u>
SD Activity Dimension	<u>Second</u>	<u>Fourth</u>	<u>First</u>	Third

Note. Any two means underlined by the same line are not significantly different, whereas any two means not underlined by the same line are significantly different. Means are arranged in descending size from left to right. All differences are at the $p < .05$ level.

Three further significant sources of variance were identified. Sequence was significant for the Evaluative dimension of the Semantic Differential [$F(3,48)=2.98$, $p<.05$] but did not attain significance over any of the other dependent measures. Two Latin Square Residuals reached significance for the TEI suggesting differential carryover effects over knowledge groups and case levels.

With the exception of the possible interaction of the LSR_{CBD} , there were no main or interaction effects for case description.

D. CASE AND TREATMENT QUESTIONNAIRE ANALYSIS

The results of the Friedman Analysis of Variance by Ranks for the responses to question three of the questionnaire are summarized in Table 21.

The results summarized in Table 21 show that, in fact, the mean ranks given the treatments differ ($\chi^2_{\text{ranks}}=46.76$). Inspection of the means show that Reinforcement and Time Out have the larger mean ranks, Medication, the smallest, with the Postive Practice mean rank falling above that for Medication. In addition, the significant Coefficient of Concordance ($\omega=.2435$) indicates that teachers agree in their ranking of preferences. The average intercorrelation between teachers is significant ($\bar{r}=.23$), again suggesting that there is some agreement in the relative ranking of treatments.

The results of pairwise comparisons between treatment rank means using Nemenyi's (1963) procedure are summarized

TABLE 21
RESULTS OF FRIEDMAN TWO-WAY ANALYSIS OF VARIANCE
ON RANKING OF TREATMENTS

Treatment	Rank Sum	Mean Rank	S.D.
Reinforcement	193.5	3.02	0.90
Time Out	191.5	3.00	0.94
Positive Practice	147.0	2.28	1.09
Medication	108.0	1.69	0.97

Friedman Test Statistic $\chi^2_{\text{ranks}} = 46.76$

$\chi^2(.05, 3) = 7.82$

Kendall Coefficient of Concordance $\omega = 0.24$

Intercorrelation between Ranks $\bar{r} = .23$

in Table 22. The results summarized in Table 22 show that while Reinforcement was ranked highest, it was not ranked significantly higher than Time Out. Both of these treatments, however, received higher rankings than the Positive Practice intervention. The Medication intervention was ranked lowest, significantly lower than any of the other three treatments. The results of the Case and Treatment Questionnaire ranking of treatments indicated that treatments were ranked differentially. Comparisons of treatment means indicated generally similar results to those found on both the TEI and Evaluative dimension of the SD.

TABLE 22
NEMENYI PAIRWISE COMPARISONS BETWEEN MEAN TREATMENT RANKS

Treatment	Mean	Difference			
		R	TO	PP	M
Reinforcement (R)	3.02	--	0.02	0.74*	1.33*
Time Out (TO)	3.00		--	0.72*	1.31*
Positive Practice (PP)	2.28			--	0.59*
Medication (M)	1.69				--

* $p < .05$

The results of the analyses presented in this chapter and possible implications are discussed in Chapter Four.

IV. SUMMARY AND CONCLUSIONS

The purpose of this chapter is to summarize the methods used and results obtained in the present study. These results are discussed in light of research findings in other investigations. The implications for further research and for education are also considered.

A. SUMMARY OF METHODS AND RESULTS

In this research, regular classroom teachers' ratings of the acceptability of four classroom treatment interventions were examined. Teachers were assigned to treatment conditions based on the score which they received on a measure of knowledge of behavioral principles as applied to children and randomly assigned within knowledge groups to one of two case descriptions. The first case described a student who is experiencing extreme emotional and behavioral problems in the classroom settings. The second case described the same child but with the insertion of additional background information suggesting a history of special class treatment. A major purpose of this research was to determine if, as hypothesized, teacher's level of knowledge of behavioral principles would significantly affect ratings of treatment acceptability. A second purpose of the study was to investigate whether teachers' ratings of treatment would reflect the different cases described. Related to these questions was the question of whether teachers would differentiate treatments in terms of acceptability. It was

hypothesized that teachers would provide significantly different ratings on the treatment acceptability measures for the four treatments presented and further that the degree of knowledge expressed on the knowledge measure would have a significant effect on the ratings teachers provide. Past research evidence supported these research questions, although research findings had not directly substantiated these specific hypotheses. The question of case differences is less clear. Research to date has been equivocal. Despite mixed evidence supporting such a hypothesis it was theorized that the background attributed to the two cases would have a significant effect on teacher ratings of treatment acceptability.

Another question addressed by the research was related to the method of measurement of acceptability ratings. The method described in the study as Treatment Acceptability Measures reflect an established methodology using four measures of treatment acceptability (Kazdin, 1980a). The present study employed, in addition to these, a questionnaire which in part asked subjects to rank the four described treatments in terms of acceptability and use. It was suggested that this type of procedure might simulate more closely the clinical decision-making process, and hypothesized that such a procedure would provide results comparable to those of the TEI treatment acceptability measure.

1. KNOWLEDGE GROUP

In general the hypothesis of knowledge group difference was supported. The high knowledge group teachers generally provided higher acceptability ratings for all treatments. This result was evident from the ratings on the TEI only. Ratings on the three dimensions of the Semantic Differential showed no significant difference for this factor.

The finding of generally higher acceptability ratings (on the TEI) for the high knowledge group lends support to the major findings reviewed in Chapter One. The majority of earlier studies had found generally more positive attitudes toward behavioral approaches to treatment with increased knowledge of behavioral techniques and underlying principles. This is a particularly interesting finding which would suggest that improved acceptability, and potential increase in use of behavioral treatments, may be facilitated through increased familiarity with the techniques and underlying principles of behavioral interventions. Presumably, the same might be said for the acceptability and knowledge of the Medication intervention. This remains an empirical question, although it may be assumed that the two knowledge groups were similarly informed regarding the use of Medication in the treatment of behavioral disorders. In a very broad sense, these findings would suggest that with increased knowledge of a variety of therapeutic approaches teachers would likely become more accepting of alternative approaches to classroom treatment. Perhaps ideal, in

settings where a variety of student needs are expressed, is a teacher willing and able to draw upon a similarly large variety of treatment approaches to meet those needs.

The finding of no significant differences between knowledge groups on the three dimensions of the Semantic Differential is interesting in light of similar findings of no difference by Burkhart, Behles, and Stumphauzer (1976). Burkhart et al. (1976) employed a semantic differential measure of attitude toward behavior modification and found no differences in attitudes for trained and control groups. Kazdin (1980a, 1980b, 1981a, in press) and Kazdin, French and Sherick (1981) have also found few significant effects on the SD other than differences in treatment type. This would suggest that while the SD provides an accurate reflection of differences between treatments, it does not appear to be a sufficiently sensitive rating procedure to consistently reflect other more subtle between groups differences.

Another comparison of interest is that of the present study with that of Kazdin, French, and Sherick (1981) which used the same measures of treatment acceptability and similar treatment descriptions. Three groups of subjects participated in the Kazdin et al. (1981) study: child inpatients, parents, and treatment staff. The authors report finding significant differences between parent and child overall ratings of treatment acceptability (i.e., TEI ratings), with parents giving higher mean ratings to

treatments than children. Treatment staff, who presumably are expert in the use of the described techniques, provided ratings which were not significantly different from either the parent or child ratings. In the present study, however, the high knowledge group rated treatments generally more positively than the low knowledge group. While the subject samples in the two studies are different, it is reasonable to suggest that the subjects in the present study are most similar to the treatment staff subjects in the Kazdin study. However, Kazdin et al. did not investigate the status of participants' knowledge of behavioral principles and did not investigate the differences between [parent, staff, child] groups in the ratings of behavioral treatments or medication making more direct comparisons of results between these two studies impossible.

2. CHILD CASE

The results of the analysis of treatment measure ratings did not reveal significant differences in the acceptability of treatments for the two cases described. There are several plausible explanations which might account for such a result. The most obvious of these is that, in fact, the cases were not different and that the additional information regarding previous special class placement provided in the case of Mike S. added no information relevant to treatment evaluation. In eight reported experiments using this methodology, Kazdin and his

colleagues (Kazdin 1980a, 1980b, 1981a, in press; Kazdin, French, & Sherrick 1981) have found acceptability rating differences between case descriptions in only one study and with manipulation of case features, such as severity of problem, found differences reflected only sporadically. In one study (1980a, Experiment 2) the TEI and Evaluative dimension of the SD showed significant effects for case description although expected differences in ratings for more and less severe cases were noted only with the TEI. Severity of case was significant on the TEI and SD Evaluative in a more recent study as well (Kazdin, in press). A major difference between the present study and earlier work by Kazdin is that the cases used in the Kazdin studies are purposely varied on a number of traditionally relevant dimensions, for example, age, sex, type of problem in order to render the results more generalizable. In contrast, the present study sought to vary only one dimension (i.e., background of class placement of the case description) in order to test the effect of prior placement information on treatment acceptability ratings in a manner more relevant to the regular classroom setting. Despite the differences in purpose, the findings of the Kazdin studies for increased acceptability in more severe cases may be quite relevant here. In the present study, in addition to the treatment acceptability procedures, teachers were asked to rate the seriousness of the problems Mike is exhibiting as part of the Case and Treatment Questionnaire. Without

exception, teachers rated these problems more than moderately serious and almost always as extremely serious, although they did indicate that this child's behavior was typical of the kinds of problems exhibited by some children in the regular classroom setting. Teachers did not differentiate the severity of the problems exhibited between case descriptions.

Also relevant here are the findings of Safran, Safran, and Orlansky (1982). In this study variations in written background information had significant effects upon subjects' perceptions of the child's behavior, and upon the degree to which children were regarded as able to function successfully in the regular classroom. This finding was true for descriptions of both hearing impaired and normal children. However this finding did not extend to the case of an acting-out child. The authors suggest that the extreme nature of the behavior relegated the written background information to a less influential position.

It is reasonable to speculate that such a result may have occurred in the present study and that any relevance that placement history may have had was masked by the perceived extreme nature of the problems the child was described as exhibiting. This explanation certainly is consistent with the findings of Keogh and Levitt (1976, cited in Safran, Safran, & Orlansky, 1982) who suggest that the salient issues are both the regular teachers knowledge of the characteristics and needs of exceptional children and

descriptive information relevant to the child's performance. A further possible, although less likely, explanation may be that teachers in the present study did not relate acceptability ratings to the presented case description.

3. TREATMENT TYPE

Results of the analysis presented in the previous chapter suggested that the regular classroom teachers sampled in the present study did differentiate treatments in terms of acceptability. Two measures of acceptability (TEI and Evaluative SD) showed that the Reinforcement intervention, a positive approach to classroom intervention for the problem behavior described, was the most acceptable. The Medication intervention was rated as the least acceptable of the four alternatives, although it was clearly seen as the most potent or powerful. That Reinforcement was seen as the least powerful is also worthy of note.

In terms of the TEI and SD Evaluative results it seems reasonable to group the interventions in various ways for the purpose of discussion. The results of analysis on these two measures clearly indicated that teachers prefer behavioral to medical interventions, particularly for the type of problems described. What is also clear, however, is that the Reinforcement intervention is seen as more acceptable than either of the two more punitive behavioral interventions, Positive Practice and Time Out. In turn both of the punitive interventions were seen as more acceptable

than use of Medication.

In general this follows the results of Kazdin (1980a, 1981, in press) and Kazdin, French, and Sherick (1981). The results of the ranking procedure reported in the analysis of the Case and Treatment Questionnaire suggest that Medication is lowest in acceptability and potential for use. Positive Practice is seen as an improvement on this. The distinction evident in ratings of Reinforcement relative to other interventions on the TEI and Evaluative SD are not so evident in the ranking procedure. Perhaps the results of the ranking procedure reflect both the acceptability as measured by the TEI and Evaluative SD as well as some contribution of the strength of the intervention or its effectiveness as reflected in the potency ratings.

Kazdin, French, and Sherick (1981), in discussion of a similar finding of high acceptability and low potency for Reinforcement concluded that parents and others need to be convinced that, in fact, Reinforcement can be very effective (i.e., potent).

The present findings also suggest that when asked about acceptability and potential use teachers consider not only features of acceptability but also the potential for the treatments under consideration to be effective in dealing with the particular problem in the described setting.

Results on the activity dimension of the SD did not reveal any significant differences between treatments in the present study. Again this finding is similar to that of

previous research. This dimension appears, in this study and to some extent in those previous, to be of limited utility in reflecting differences between treatments.

Two additional comments should be made with regard to the treatment acceptability findings. The first of these is that while Medication was given the lowest mean rating on the TEI (50.97) and the SD Evaluative (17.5) these results still suggest moderate overall acceptability, as a score of 52.5 on the TEI and 17.5 on the SD Evaluative represent neutral ratings. In other words, teachers have not suggested outright that Medication is an unacceptable intervention. Similarly, the mean ratings for Reinforcement on the TEI (79.03) and SD Evaluative (27.05) while quite favorable, do not reflect outright acceptance as these values represent 75% and 77% of the maximum value for their respective instrument. A second important observation to be made with regard to the relative acceptability of treatments is that it is likely that many of the teachers sampled in the present study use, daily, variations of several of the described treatment interventions. None of these interventions represent novel approaches to management of problem classroom behavior. Likely the most unfamiliar of these is the Medication intervention and in particular the specific drug treatment described.

B. IMPLICATIONS FOR EDUCATION

At the present time the teaching mandate in regular school classrooms includes an expectation that teachers take on the major responsibility for making changes in student behavior not only with regard to the student's academic performance but also in myriad other ways. Educational and legal policy, as well as professional ethics dictate that efforts directed toward such change should occur despite such adverse conditions as intellectual impairment and emotional and behavioral disturbance. Given these demands teachers increasingly have need of effective approaches which will aid in the carrying out of this mandate.

Despite the wealth of literature which exists substantiating the effectiveness of particular procedures and indeed of several general approaches in the classroom setting such procedures and approaches have not received anywhere near universal acceptance and use. The present study suggests that one factor which may explain this situation is a lack of knowledge of the procedures and underlying principles. A further indication of note here is the finding that those teachers who were in the high knowledge group generally indicated having specific training in child management, whereas those in the low knowledge group indicated not having such training. Typically teacher training programs stress curriculum and methodology but do not routinely prepare teachers to be effective managers and behavior change agents. This becomes particularly crucial in

the case of those students who are not capable of self management in the classroom setting. Teachers must be well informed and capable in both general management and behavior change skills as well as in the understanding and application of the underlying principles of learning and behavior. The obvious implication here is that both pre-service and inservice teacher training must begin to devote a greater emphasis to development and application of relevant knowledge and skills. The work of several authors (Merrett & Wheldall, 1982; Mohlman, Coladarci & Gage, 1982; Robinson & Swanton, 1980;) certainly supports such a contention.

A second implication which follows from this research and from the findings of Witt, Martens, and Elliot (1984) is that teachers often are working in less than ideal settings and under constraints of time and personnel which seem to dictate use of procedures which may not be the best but which are pragmatic. Ford and Kendall (1979) suggest that even those clinicians committed to a particular orientation do not necessarily bypass more pragmatic solutions in favor of theoretically consistent procedures. In attempts at short term change and intervention, consultants must be aware that treatment recommendations will vary in their success not only as a result of the verified effectiveness of the procedure but also of the acceptability for use by the particular teacher in the given settings. Kazdin (Kazdin, French, & Sherick, 1981) has suggested that available

treatments may be presented and delivered in ways which make the treatments "more palatable to persons who, for whatever reasons, are not likely to adhere to treatment procedures" (p. 906). Witt and Elliot support this point of view and further suggest that interventions be described from the perspective of the attitudes and values of the teacher rather than from a particular theoretical perspective, accepting the fact that in short consultation sessions it is probably impossible to alter markedly a teacher's beliefs about the causes of behavior. Thus the implications stated above are related in that the acceptability of an intervention is a likely prerequisite for its eventual use by a classroom teacher. This study would suggest that increased knowledge of behavioral principles facilitates greater acceptability of effective behavioral techniques. Acceptability is seen as a necessary condition for use--which is the primary objective. Given that individuals do vary in their knowledge and theoretical orientation, acceptability of treatments can be enhanced through appropriate presentation and application. An approach which is both acceptable and effective is likely to be used in the future. Similarly, an approach which is not seen as acceptable is not likely to be used at all.

C. IMPLICATIONS FOR FUTURE RESEARCH

A major goal of the study of treatment acceptability is to build a body of research which will facilitate the identification and use of effective and acceptable treatments. Witt and Elliot (in press) have suggested that the dimensions of treatment acceptability, integrity, effectiveness, and use are reciprocally and sequentially interrelated (cf. Yeaton & Sechrest, 1981). The ultimate test of a treatment is its use. Witt and Elliot (in press) suggest that the initial issue in treatment selection is acceptability, and that use then is related to how effective the procedure proves to be. The integrity with which treatment is applied often determines its effectiveness. Future research should focus on building strong evidence relating each of these dimensions. An obvious and important area for immediate investigation is that of the link between acceptability and use.

The results of the ranking procedure in the present study may be viewed as encouraging in terms of furthering the development of relevant procedures for the measurement of treatment acceptability. This procedure was proposed as an initial attempt at approximating the clinical decision making process. Future research may extend such an approach to sample more of the dimensions of interest. One possible application could involve ranking of several treatments on each of the areas sampled by questions on the TEI. For example subjects could be asked to examine several

treatments and to rank each on its relative effectiveness or how humane each treatment is likely to be.

In the field of special education in particular and education in general it is clear that it is often easier to find 'good assessments and diagnoses' than it is to find good treatment. Similarly it is often easier to find good theory than it is to find good application, particularly with respect to dealing with children's problem behavior in the regular classroom setting. There is a need for research which assesses the implementation of actual treatment recommendations. Related to this are issues of acceptability but also of 'integrity' of methods (i.e., do people do what they say they do), effectiveness, and the ability and willingness of teachers and other school and treatment personnel to carry out recommended treatments. One example of such research might involve follow-up on implementation of treatment recommendations from a diagnostic center.

There is the suggestion, from Witt and Elliot (in press), that in addition to past experience with changing the targeted problem, one's philosophical orientation or general approach to changing behavior will affect initial judgements about the acceptability of treatments. With regard to the issue of philosophical orientation, the work of Miller (1981) and Norcross & Prochaska (1983) represent attempts to characterize in a reliable manner teachers' affiliation to various orientations and theories. Such approaches may provide important dimensions related to the

acceptability of interventions from a variety of theoretical orientations.

A further implication presented is in response to the needs identified by Kiesler (1983). Future research in the areas of program effectiveness and consumer evaluation of treatments will need to establish what Kiesler calls an 'attitude pretest' in order to better understand the impact of treatments acceptability in such research settings. The procedures described in this study and previously by Kazdin may be most useful in providing such data.

Finally, further research is needed which assesses the relative acceptability of treatments among significant groups involved in the intervention enterprise--the clinicians and other consultants who recommend treatment, the teachers and parents who implement treatments, and the children who are the object of treatment regimes. Present research suggests disparate views, both between groups of individuals and between alternative treatments.

REFERENCES

- Adelman, H., Taylor, L., Fuller, W. & Nelson, P. (1979). Discrepancies among student, parent, and teacher ratings of the severity of a student's problems. American Educational Research Journal, 16(1), 38-41.
- Baer, D. M. & Bushell, D., Jr. (1981). The future of behavior analysis in the schools? Consider its recent past, and then ask a different question. School Psychology Review, 10(2), 259-270.
- Baer, G. T., Goodall, R., & Brown, L. (1983). Discipline in the classroom: Perceptions of middle grade teachers. The Clearinghouse, 57(3), 139-142.
- Bandura, A. (1975). The ethics and social purpose of behavior modification. Annual Review of Behavior Therapy, 3, 13-20.
- Bandura, A. (1977). Social learning theory. Englewood Cliffs, N. J.: Prentice-Hall.
- Barling, J. & Wainstein, T. (1979). Attitudes, labeling bias, and behavior modification in work organizations. Behavior Therapy, 10(1), 129-136.
- Barlow, D. H. (1980). Behavior therapy: The next decade. Behavior Therapy, 11(3), 315-328.
- Barnes, S. E., Fors, S. W., & Decker, W. H. (1980). Ethical issues in health education. Health Education, 11(2), 7-9.
- Becker, W. C. (1971). Parents as teachers. Champaign, IL: Research Press.
- Benassi, V. & Lanson, R. (1972). A survey of the teaching of behavior modification in colleges and universities. American Psychologist, 27, 1063-1069.
- Berger, M. (1979). Behavior modification in education and professional practice: The dangers of a mindless technology. Bulletin of the British Psychological Society, 32, 418-419.
- Bergin, A. E. (1980). Behavior therapy and ethical relativism: Time for clarity. Journal of Consulting and Clinical Psychology, 48(1), 11-13.
- Berman, L. (1927). The religion called behaviorism. New York: Boni & Liveright.

- Bornstein, P. H. & Rychtarik, R. G. (1983). Consumer satisfaction in adult behavior therapy: Procedures, problems, and future perspectives. Behavior Therapy, 14, 191-208.
- Boutillier, J. (1981). Behavior modification works. Atlantic Canada Teacher, 10, 9-13.
- Brophy, J. E. & Good, T. L. (1974). Teacher-student relationships: Causes and consequences. New York: Holt, Rinehart, & Winston.
- Bruning, J. L. & Kintz, B. L. (1977). Computational handbook of statistics (2nd ed.). Glenview, IL: Scott, Foresman.
- Burkhart, B. R., Behles, M. W., & Stumphauzer, J. S. (1976). Training juvenile probation officers in behavior modification: Knowledge, attitude change, or behavioral competence? Behavior Therapy, 7, 47-53.
- Burton, J. K. (1981). Behavioral technology: Foundation for the future. Educational Technology, 21, 21-28.
- Calhoun, K. S. & Turner, S. M. (1981). Historical perspectives and current issues in behavior therapy. In S. M. Turner, K. S. Calhoun, & H. Adams (Eds.), Handbook of clinical behavior therapy (pp. 1-11). New York: Wiley.
- Camplese, D. A., O'Brube, W. S., & Hale, R. E. (1979). The relationship between teaching experience and attitude toward behaviorally stated objectives. College Student Journal, 13(3), 221-223.
- Carey, R. G. & Bucher, B. (1983). Positive practice overcorrection: The effects of duration of positive practice on acquisition and response reduction. Journal of Applied Behavior Analysis, 16(1), 101-109.
- Carrera, F., III, & Adams, P. L. (1970). An ethical perspective on operant conditioning. Journal of the American Academy of Child Psychiatry, 9, 607-623.
- Christian, W. P., Clark, H. B., & Luke, D. E. (1981). Client rights in clinical counseling services for children. In G. T. Hannah, W. P. Christian, & H. B. Clark (Eds.), Preservation of clients rights: A handbook for practitioners providing therapeutic, educational and rehabilitative services (pp. 19-41). Riverside, CA: Free Press.

- Cochran, W. G. & Cox, G. M. (1950). Experimental designs. New York: Wiley.
- Dipasquale, G. W. (1980). Legal issues in using behavior modification in schools: A comment on James and Allon. Canadian Psychologist, 21(3), 132-134.
- Doke, L. A. & Flippo, J. R. (1981). Behavior therapy in the treatment of childhood disorders. In L. Michelson, M. Hersen, & S. M. Turner (Eds.), Future perspectives in behavior therapy (pp. 21-43). New York: Plenum.
- Dubno, P., Hillburn, D., Robinson, G., Sandler, D., Trani, J., & Weingarten, E. (1978). An attitude toward behavior modification scale. Behavior Therapy, 9, 99-108.
- Eber, M. & O'Brien, J. M. (1982). Psychotherapy in the movies. Psychotherapy: Theory, Research and Practice, 19(1), 116-120.
- Elliot, S.N., Witt, J.C., & Galvin, G. (1983). Children's judgements about the acceptability of traditional and behavioral classroom interventions. Manuscript submitted for publication.
- Engelhardt, H. T. & McCullough, L. B. (1982). Ethics in psychiatry. In S. Arieti & H. K. H. Brodick (Eds.), American handbook of psychiatry: Vol. 7. Advances and new directions (2nd ed.) (pp. 795-818). New York: Basic Books.
- Eysenck, H. J. (1971). Behavior therapy as a scientific discipline. Journal of Consulting and Clinical Psychology, 36, 314-319.
- Eysenck, H. J. (1970). Behavior therapy and its critics. Journal of Behavior Therapy and Experimental Psychiatry, 1, 5-15.
- Eysenck, H. J. & Beech, H. R. (1971). Counterconditioning and related methods. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change (pp. 60-97). New York: Wiley.
- Farnham-Diggory, S. (1981). But how do we shape up rigorous behavioral analysts? Developmental Review, 1(1), 58-60.
- Feldman, M. P. & Peay, J. (1982). Ethical and legal issues. In A. S. Bellack, M. Hersen, & A. E. Kazdin (Eds.), International handbook of behavior modification and therapy (pp. 231-261). New York: Plenum.

- Ford, J. D. & Kendall, P. C. (1979). Behavior therapists' professional behaviors: A survey study. Professional Psychology, 10(6), 772-773.
- Foster, B. & Ysseldyke, J. (1976). Expectancy and halo effects as a result of artificially induced teacher bias. Contemporary Educational Psychology, 2, 37-45.
- Foxx, R. M. & Jones, J. R. (1978). A remediation program for increasing the spelling achievement of elementary and junior high school students. Behavior Modification, 2(2), 211-230.
- Foxx, R. M. & Shapiro, S. T. (1978). The timeout ribbon: A nonexclusionary timeout procedure. Journal of Applied Behavioral Analysis, 11, 125-136.
- Frey, S. H. (1970). Teachers and behavioral modification. Phi Delta Kappan, 55, 634-635.
- Friedman P. R. (1975). Legal regulation of applied behavior analysis in mental institutions and prisons. Arizona Law Review, 17, 39-104.
- Furtkamp, E., Giffort, D., & Schiers, W. (1982). In-class evaluation of behavior modification knowledge: Parallel tests for use in applied settings. Journal of Behavior Therapy and Experimental Psychiatry, 13, 131-134.
- Garfield, S. L. (1983). Some comments on consumer satisfaction in behavior therapy. Behavior Therapy, 14, 237-241.
- Gargiulo, R. M. & Yonker, R. J. (1983). Assessing teachers' attitude toward the handicapped: A methodological investigation. Psychology in the Schools, 20, 229-233.
- Giles, T. R. (1983). Probable superiority of behavioral interventions-I: Traditional comparative outcomes. Journal of Behavior Therapy and Experimental Psychiatry, 14, 29-32.
- Goldiamond, I. (1975a). Singling out behavior modification for legal regulation: Some effects on patient care, psychotherapy, and research in general. Arizona Law Review, 17, 105-126.
- Goldiamond, I. (1975b). Toward a constructional approach to social problems: Ethical and constitutional issues raised by applied behavior analysis. Annual Review of Behavior Therapy, 3, 21-63.

- Guba, E. .(1972). The failure of educational evaluation. In C. Weiss (Ed.), Evaluation action programs, (pp. 250-266). Boston: Allyn & Bacon.
- Gutkin, T. B., Singer, J. H., & Brown, R. (1980). Teacher reactions to school-based consultation services: A multivariate analysis. Journal of School Psychology, 18(2), 126-134.
- Guttmann, J. (1982). Pupils', teachers', and parents' causal attributions for problem behavior at school. Journal of Educational Research, 76(1), 14-21.
- Hall, R. V. (1971). Managing behavior (vol. 2). Lawrence, KS: H. & H. Enterprises.
- Hannah, M. E. & Pliner, S. (1983). Teacher attitudes toward handicapped children: A review and synthesis. School Psychology Review, 12(1), 12-25.
- Hargreaves, W. A. & Attkisson, C. C. (1978). Evaluating program outcomes. In C. C. Attkisson, W. A. Hargreaves, & J. E. Sorensen (Eds.), Evaluation of human service programs (pp. 20-37). New York: Academic Press.
- Haring, N. G. (1957). A study of the attitudes of classroom teachers toward exceptional children. Dissertation Abstracts, 17, 103-104. (University Microfilms No. 00-19, 378).
- Heldman, A. W. (1973). Social psychology versus the First Amendment freedoms, due-process liberty, and limited government. Cumberland-Samford Law Review, 4, 1-40.
- Hickey, D. F. (1977). School counselors' attitudes toward behavior modification. Personnel and Guidance Journal, 55(8), 477-480.
- Hobbs, S. A., Walle, D. L., & Caldwell, H. S. (1984). Maternal evaluation of social reinforcement and time-out: Effects of brief parent training. Journal of Consulting and Clinical Psychology, 52(1), 135-136.
- Hochbaum, G. (1980). Ethical dilemmas in health education. Health Education, 11(2), 4-6.
- Holland, J. G. (1975). Behavior modification for prisoners, patients and other people as a prescription for the planned society. Mexican Journal of Behavior Analysis, 1(1), 81-95.

- Hughes, S. L., Kauffman, J. M., & Wallace, G. (1973). What do labels really mean to classroom teachers? Academic Therapy, 8, 285-289.
- Innis, N. K. (1981). Ethics of behavior modification: Behavioral and medical psychology. Social Science and Medicine: Part. F. Medical and Social Ethics, 15(1), 69-73.
- James, E. L. & Allon, R. (1979). Legal issues and behavior modification: A Canadian perspective. Canadian Psychological Review, 20(4), 175-183.
- Jeger, A. M. & McClure, G. (1979). Attitudinal effects of undergraduate behavioral training. Teaching of Psychology, 6(4), 226-228.
- Jeger, A. M. & McClure, G. (1982). An experimental evaluation and process analysis of a behavioral consultation program. In A. M. Jeger & R. S. Slotnick (Eds.), Community mental health and behavioral ecology: A handbook of theory, research & practice, (pp. 389-401). New York: Plenum.
- Jones, R. T. & Kazdin, A. E. (1981). Childhood behavior problems in school. In S. M. Turner, K. S. Calhoun, & H. E. Adams (Eds.), Handbook of clinical behavior therapy (pp. 568-606). New York: Wiley.
- Kazdin, A. E. (1979). Fictions, factions and functions of behavior therapy. Behavior Therapy, 10, 629-654.
- Kazdin, A. E. (1980a). Acceptability of alternative treatments for deviant child behavior. Journal of Applied Behavior Analysis, 13, 259-273.
- Kazdin, A. E. (1980b). Acceptability of time out from reinforcement procedures for disruptive child behavior. Behavior Therapy, 11, 329-344.
- Kazdin, A. E. (1981a). Acceptability of child treatment techniques: The influence of treatment efficacy and adverse side effects. Behavior Therapy, 12, 493-506.
- Kazdin, A. E. (1981b). Behavior modification in education: Contributions and limitations. Developmental Review, 1(1), 34-57.
- Kazdin, A. E. (1981c). Uses and abuses of behavior modification in education: A rejoinder. Developmental Review, 1(1), 61-62.

- Kazdin, A. E. (1982). The token economy a decade later. Journal of Applied Behavior Analysis, 15, 431-445.
- Kazdin, A. E. (1983). Psychiatric diagnosis, dimensions of dysfunction, and child behavior therapy. Behavior Therapy, 14, 73-99.
- Kazdin, A. E. (in press). Acceptability of aversive procedures and medication as treatment alternatives for deviant child behavior. Journal of Abnormal Child Psychology.
- Kazdin, A. E. & Cole, P. M. (1981). Attitudes and labeling biases toward behavior modification: The effects of labels, content, and jargon. Behavior Therapy, 12, 56-68.
- Kazdin, A. E., French, N. H., & Sherick, R. B. (1981). Acceptability of alternative treatments for children: Evaluations by inpatient children, parents and staff. Journal of Consulting and Clinical Psychology, 49, 900-907.
- Kazdin, A. E. & Hersen, M. (1980). The current status of behavior therapy. Behavior Modification, 4(3), 283-302.
- Kazdin, A. E. & Wilson, G. T. (1978). Evaluation of behavior therapy: Issues, evidence and new research strategies. Cambridge, MA: Ballinger.
- Kelly, H. H. (1950). The warm-cold variable in first impressions of persons. Journal of Personality, 18, 431-439.
- Kent, R. N. & O'Leary, K. D. (1976). A controlled evaluation of behavior modification with conduct problem children. Journal of Consulting and Clinical Psychology, 44, 586-596.
- Keogh, B. K. & Levitt, M. L. (1976). Special education in the mainstream: A confrontation of limitations? Focus on Exceptional Children, 8(1), 1-11.
- Kerr, M. M. & Lambert, D. L. (1982). Behavior modification of children's written language. Progress in Behavior Modification, 13, 79-108.
- Kiesler, C. A. (1983). Social psychological issues in studying consumer satisfaction with behavior therapy. Behavior Therapy, 14, 226-236.

- Kirigin, K. A., Braukmann, C. J., Atwater, J. D., & Wolf, M. M. (1982). An evaluation of teaching family (Achievement Place) group homes for juvenile offenders. Journal of Applied Behavior Analysis, 15, 1-16.
- Kitchener, R. F. (1980a). Ethical relativism and behavior therapy. Journal of Consulting and Clinical Psychology, 48, 1-7.
- Kitchener, R. F. (1980b). Ethical relativism, ethical naturalism, and behavior therapy. Journal of Consulting and Clinical Psychology, 48, 14-16.
- Kitchener, R. F. (1980c). Ethical skepticism and behavior therapy: A reply to Ward. Journal of Consulting and Clinical Psychology, 48, 649-651.
- Knapp, J. R. & Delprato, D. J. (1980). Willpower, behavior therapy and the public. Psychological Record, 30, 477-482.
- Krasner, L. (1976). On the death of behavior modification: Some comments from a mourner. American Psychologist, 31, 387-388.
- Lahey, B. B., Delamater, A., & Kupfer, D. (1981). Intervention strategies with hyperactive and learning-disabled children. In S. M. Turner, K. S. Calhoun, & H. E. Adams (Eds.), Handbook of clinical behavior therapy (pp. 607-634). New York: Wiley.
- Lahey, B. B. & Rubinoff, A. (1981). Behavior therapy in education. In L. Michelson, M. Hersen, & S. M. Turner (Eds.), Future perspectives in behavior therapy (pp. 27-43). New York: Plenum.
- Lambert, N. M. (1976). Children's problems and classroom interventions from the perspective of classroom teachers. Professional Psychology, 7, 504-517.
- Lazarus, A. A. (1977). Has behavior therapy outlived its usefulness? American Psychologist, 32, 550-554.
- Lebow, J. (1982). Consumer satisfaction with mental health treatment. Psychological Bulletin, 91, 244-259.
- Lewin, P., Nelson, R. E., & Tollefson, N. (1983). Teacher attitudes toward disruptive children. Elementary School Guidance and Counseling, 17(3), 188-193.
- Lindsey, J. D. & Frith, G. H. (1983). Detrimental student behaviors as perceived by elementary teachers. Education, 103(3), 306-308.

- Luiselli, J. K. (1981). Consultation in the residential treatment of visually impaired children. Journal of Visual Impairment and Blindness, 75(9), 353-358.
- Maher, B. A. (1978). Stimulus sampling in clinical research: Representative design reviewed. Journal of Consulting and Clinical Psychology, 46(4), 643-647.
- Mannarino, A. P. & Durlak, J. A. (1980). Implementation and evaluation of service programs in community settings. Professional Psychology, 11(2), 220-227.
- Marholin, D., Taylor, R. L., & Warren, S. A. (1978). Learning to apply psychology: Didactic training experience and opinions about behavior modification. Teaching of Psychology, 5(1), 23-26.
- Marotz, B. (1983, April). Alternatives to behavioral classroom management. Paper presented at the meeting of the Council of Exceptional Children. Detroit, MI.
- Martin, G. L. (1981). Behavior modification in Canada in the 1970's. Canadian Psychology, 22(1), 7-22.
- Mazza, J. & Pumroy, D. K. (1975). A review of evaluation of behavior modification programs. Psychological Record, 25, 111-121.
- McMahon, R. J. & Forehand, R. L. (1983). Consumer satisfaction in behavioral treatment of children: Types, issues and recommendations. Behavior Therapy, 14, 209-225.
- McMahon, R. J., Forehand, R., & Griest, D. L. (1981). Effects of knowledge of social learning principles on enhancing treatment outcome and generalization in a parent training program. Journal of Consulting and Clinical Psychology, 49(4), 526-532.
- Merrett, F. & Wheldall, K. (1978). Playing the game: A behavioral approach to classroom management in the junior school. Educational Review, 30(1), 41-50.
- Merrett, F. & Wheldall, K. (1982). Does teaching student-teachers about behavior modification techniques improve their teaching performance in the classroom? Journal of Education for Teaching, 8(1), 67-75.
- Michael, J. (1980). Flight from behavior analysis: Presidential address, ABA 1980. Behavior Analyst, 3(2), 1-22.

- Miller, P. J. (1981). Perspectives, teachers' opinions and the ideational basis of education. Alberta Journal of Educational Research, 27(2), 182-191.
- Milne, D. C. (1982). A comparison of two methods of teaching behavior modification to mental handicap nurses. Behavioral Psychotherapy, 10(1), 54-64.
- Mohlman, G. G., Coladarci, J., & Gage, N. L. (1982). Comprehension and attitude as predictors of implementation of teacher training. Journal of Teacher Education, 33(1), 31-36.
- Molloy, G. N. (1980). Training parents of retarded children as behavior therapists: What we need to know. Mental Retardation Bulletin, 8(1), 3-12.
- Musgrove, W. J. (1974). A scale to measure attitudes toward behavior modification. Psychology in the Schools, 11, 392-396.
- Musgrove, W. J. & Harms, R. A. (1975). Teachers' attitudes toward behavior modification. Journal of Student Personnel Association for Teacher Education, 13, 133-138.
- Myers, J. L. (1979). Fundamentals of experimental design (3rd ed.). Boston: Allyn and Bacon.
- Natrella, M. G. (1963). Experimental statistics. Washington, DC: U.S. Government Printing Office.
- Necco, E. G. (1970). The effects of labeling on teachers' perceptions (Doctoral dissertation, University of Virginia 1970). Dissertation Abstracts International, 31, 4587A (University Microfilms No. 70-26, 627).
- Nelson, L. (1974). Guide to Lertap use and interpretation. Dunedin, New Zealand: University of Otago..
- Nemenyi, P. (1963). Distribution-free multiple comparisons. Unpublished doctoral thesis. Princeton: Princeton University.
- Norcross, J. C. & Prochaska, J. O. (1983). Clinician's theoretical orientations: Selection, utilization, and efficacy. Professional Psychology: Research and Practice, 14(2), 197-208.

- Norton, G. R., Austen, S., Allen, G. E., & Hilton, J. (1983). Acceptability of time out from reinforcement procedures for disruptive child behavior: A further analysis. Child and Family Behavior Therapy, 5(2), 31-41.
- Nunnally, J. M., Jr. (1970). Introduction to psychological measurement. Toronto: McGraw-Hill.
- O'Dell S. L., Tarler-Benlolo, L., & Flynn, J. M. (1979). An instrument to measure knowledge of behavioral principles as applied to children. Journal of Behavior Therapy and Experimental Psychiatry, 10(1), 29-34.
- O'Leary, K. D. (1984). The image of behavior therapy: It is time to take a stand. Behavior Therapy, 15, 219-233.
- O'Leary, S. G. & O'Leary, K. D. (1977). Ethical issues of behavior modification research in schools. Psychology in the Schools, 14(3), 299-307.
- Ollendick, T. H., Matson, J. L., Esveltd-Dawson, K., & Shapiro, E. S. (1980). Increasing spelling achievement: An analysis of treatment procedures utilizing an alternative treatments design. Journal of Applied Behavior Analysis, 13, 645-654.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). The measurement of meaning. Urbana, IL: University of Illinois Press.
- Patterson, G. R. (1971). Families. Champaign, IL: Research Press.
- Patterson, G. R. & Gullion, M. E. (1968). Living with children. Champaign, IL: Research Press.
- Petrie, P., Brown, K., Piersal, W. C., Frinfrock, S. R., Schelble, M., Leblanc, C. P., & Kratochwill, T. R. (1980). The school psychologist as behavioral ecologist. Journal of School Psychology, 18(3), 222-233.
- Pevsner, R. (1982). Group parent training versus individual family therapy: An outcome study. Journal of Behavior Therapy and Experimental Psychiatry, 13(2), 119-122.
- Phillips, L. W. (1978). The soft underbelly of behavior therapy: Pop behavior modification. Journal of Behavior Therapy and Experimental Psychiatry, 9, 139-140.

- Piersel, W. C. & Gutkin, T. B. (1983). Resistance to school-based consultation: A behavioral analysis of the problem. Psychology in the Schools, 20, 311-320.
- Pollack, M. J., & Sulzer-Azaroff, B. (1981). Protecting the educational rights of the handicapped child. In G. T. Hannah, W. P. Christian, & H. B. Clark (Eds.), Preservation of clients rights: A handbook for practitioners providing therapeutic, educational and rehabilitative services (pp. 61-82). Riverside, CA: Free Press.
- Reese, S. C., Murphy, R. J., & Filipczak, J. (1981). Assessment of multiple behavioral procedures on academic and social classroom behavior. Psychology in the Schools, 18, 349-355.
- Repucci, N. D. (1977). Implementation issues for the behavior modifier as institutional change agent. Behavior Therapy, 8, 594-605.
- Richardson, M. W. & Kuder, G. F. (1939). The calculation of test reliability coefficients based upon the method of rational equivalence. Journal of Educational Psychology, 30, 681-687.
- Robinson, V. & Swanton, C. (1980). Generalization of behavioral teacher training. Review of Educational Research, 50(3), 486-496.
- Roos, P. (1974). Human rights and behavior modification. Mental Retardation, 12(3), 3-6.
- Rosenbaum, A., O'Leary, K. D., & Jacob, R. G. (1975). Behavioral intervention with hyperactive children: Group consequences as a supplement to individual contingencies. Behavior Therapy, 6, 315-323.
- Ryan, B. A. (1976). Teacher attitudes toward behavior modification: One year after an in-service training program. Behavior Therapy, 7, 264-265.
- Safran, S. P., Safran, J. S., & Orlansky, M. D. (1982). The effects of prior information upon perceptions of exceptional children. Contemporary Educational Psychology, 7, 384-394.
- Sapon-Shevin, M. (1982). Ethical issues in parent training programs. Journal of Special Education, 16(3), 341-357.
- Saunders, J. T. (1975). The social identity of behavior modification. Unpublished doctoral dissertation, Yale University.

- Saunders, J. T. & Reppucci, N. D. (1978). The social identity of behavior modification. Progress in Behavior Modification, 6, 143-158.
- Seaton, H. W. & Aaron, R. L. (1978). Modification of pupil attitude toward reading through positive reinforcement scheduling. Reading Improvement, 15(2), 96-100.
- Sheldon-Wildgin, J. & Risley, T. R. (1982). Balancing clients rights: The establishment of human rights and peer review committees. In A. S. Bellack, M. Hersen, & A. E. Kazdin (Eds.), International handbook of behavior modification and therapy (pp. 263-289). New York: Plenum.
- Sidman, M. (1960). Tactics of scientific research: Evaluating experimental data in psychology. New York: Basic Books.
- Skinner, B. F. (1971). Beyond freedom and dignity. New York: Knopf.
- Smith, D. K. (1983). Classroom management: Approaches of regular education teachers and learning disabled resource teachers. Psychology in the Schools, 20, 363-366.
- Spece, R. G., Jr. (1972). Conditioning and other technologies used to 'Treat?' 'Rehabilitate?' 'Demolish?' prisoners and mental patients. Southern California Law Review, 45, 616-684.
- Stainbrook, G. & Green, L. W. (1982). Behavior and behaviorism in health education. Health Education, 13, 14-19.
- Stolz, S. B. (1978a). Ethical issues in behavior modification. San Francisco: Jossey-Bass.
- Stolz, S. B. (1978b). Ethics of social and educational interventions: Historical context and a behavioral analysis. In T. A. Brigham & A. C. Catania (Eds.), Handbook of applied behavior analysis: Social and instructional processes (pp. 652-676). New York: Wiley.
- Stolz, S. B. (1981). Adoption of innovations from applied behavioral research: "Does anybody care?". Journal of Applied Behavior Analysis, 12, 491-505.
- Stouffer, G. A. W., Jr., & Owen, J. (1955). Behavior problems of children as identified by today's teachers and compared with those reported by E. K. Wickman. Journal of Educational Research, 48, 321-331.

- Stumphauzer, J. S. (1981). Behavioral approaches to juvenile delinquents: Future perspectives. In L. Michelson, M. Hersen, & S. M. Turner (Eds.), Future perspectives in behavior therapy (pp. 65-80). New York: Plenum.
- Szasz, T. S. (1978). Behavior therapy: A critical review of the moral dimension of behavior modification. Journal of Behavior Therapy and Experimental Psychiatry, 9(3), 199-203.
- Throll, D. & Ryan, B. A. (1976). Research note: Teacher attitudes toward behavior modification. New Zealand Journal of Educational Studies, 11, 68-71.
- Treiber, F. A. & Lahey, B. B. (1983). Toward a behavioral model of academic remediation with learning disabled children. Journal of Learning Disabilities, 16(2), 111-116.
- Turkat, I. D. & Feuerstein, M. (1978). Behavior modification and the public misconception. American Psychologist, 33, 194.
- Turkat, I. D. & Forehand, R. (1980). Critical issues in behavior therapy. Behavior Modification, 4(4), 445-464.
- Turkat, I. D., Harris, F. G., & Forehand, R. (1979). Assessment of the public reaction to behavior modification. Journal of Behavior Therapy and Experimental Psychiatry, 10(2), 101-103.
- Tyler, W. J. (1982). Behaviorism and schooling: Issues and ideas in education. Journal of Social Policy, 11, 285.
- Vane, J. R. (1972). A school behavior modification program: Teacher attitudes a year later. Behavior Therapy, 3, 41-44.
- Wells, K. C. & Forehand, R. (1981). Childhood behavior problems in the home. In S. M. Turner, K. S. Calhoun, & H. E. Adams (Eds.), Handbook of clinical behavior therapy (pp. 12-24). New York: Wiley.
- Wheldall, K. & Congreve, S. (1980a). Attitudes of British teachers toward behavior modification. Educational Review, 32(1), 53-65.
- Wheldall, K. & Congreve, S. (1980b). Teachers and behavior modification: What do they think of it so far? British Journal of Educational Psychology, 50, 92-93.

- Wherry, J. N. (1983). Some legal considerations and implications for the use of behavior modification in the schools. Psychology in the Schools, 20(1), 46-51.
- Wickman, G. K. (1938). Teachers and behavior problems. New York: Commonwealth Fund. New York: Plenum.
- Wilson, G. T. (1982). Psychotherapy process and procedure: The behavioral mandate. Behavior Therapy, 13(3), 291-312.
- Winer, B. J. (1962). Statistical principles in experimental design. New York: McGraw-Hill.
- Witt, J. C. (1983). The relative acceptability of teacher--implemented versus principal--implemented interventions. Unpublished manuscript.
- Witt, J. C., & Elliott, S. N. (in press). Acceptability of classroom management strategies. In T. R. Kratochwill (Ed.), Advances in school psychology. Hillside, NJ: Lawrence Erlbaum & Associates.
- Witt, J. C., Elliot, S. N., & Martens, B. K. (in press). Acceptability of behavioral interventions used in classrooms: The influence of teacher time, severity of behavior problems, and type of intervention. Behavioral Disorders.
- Witt, J. C., Martens, B. K., & Elliot, S. N. (1984). Factors affecting teachers' judgement of the acceptability of behavioral interventions: Time involvement, behavior problem severity, and type of intervention. Behavior Therapy, 15(2), 204-209.
- Witt, J. C., Moe, G., Gutkin, T. B., & Andrews, L. (in press). The effect of saying the same thing in different ways: The problem of language and jargon in school-based consultation. Journal of School Psychology.
- Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. Journal of Applied Behavior Analysis, 11(2), 203-214.
- Wollersheim, J. P., McFall, M. E., Hamilton, S. B., Hickey, C. S., & Bordewick, M. C. (1980). Effects of treatment rationale and problem severity on perceptions of psychological problems and counseling approaches. Journal of Counseling Psychology, 27(3), 225-231.

- Wolpe, J. (1978). The humanity of behavior therapy. Journal of Behavior Therapy and Experimental Psychiatry, 9(3), 205-209.
- Woods, T. S. (1982). Procedures and policies for regulating field-based behavioral technology. Behavioral Disorders, 8(1), 32-40.
- Woolfolk, A. E., Woolfolk, R. L., & Wilson, G. T. (1977). A rose by any other name . . .: Labeling bias and attitudes toward behavior modification. Journal of Consulting and Clinical Psychology, 45(2), 184-191.
- Woolfolk, R. L. & Woolfolk, A. E. (1979). Modifying the effect of the behavior modification label. Behavior Therapy, 10(4), 575-578.
- Wray, L. D. & Steer, M. (1980). The implementation of behaviour modification programs in non-clinical settings. Mental Retardation Bulletin, 8(1), 13-26.
- Yeaton, W. H. & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatment: Strength, integrity and effectiveness. Journal of Consulting and Clinical Psychology, 49, 156-167.
- Young, L. D. & Patterson, J. N. (1981). Information and opinions about behavior modification. Journal of Behavior Therapy and Experimental Psychiatry, 12(3), 189-196.

APPENDIX A

Introduction, Consent, and Demographic Data Forms

PROJECT TITLE:

Teacher Ratings of Alternative Classroom Treatment Strategies.

PROJECT RESEARCHER:

William T. McKee, under the direction of Dr. Barry Munro, Professor, Department of Educational Psychology and Special Education, University of British Columbia.

PURPOSE AND PROCEDURES:

Individuals participating in this project will be evaluating several alternative classroom treatment strategies for dealing with children's deviant classroom behavior. Participants will be asked to complete an initial sixteen-item multiple choice questionnaire on approaches to dealing with some difficult but typical child behavior. Participants will then be presented with an evaluation package which will be used in the ratings procedure. Participants will be asked to complete two ratings forms following each of four presentations of treatment alternatives. The entire procedure should require no more than 45 minutes.

Participants will be asked to provide some minimal personal data (e.g. gender, age, highest level of education attained). In order to maintain complete individual

confidentiality and to ensure candid responses, the information, opinions, and data presented by all subjects will be coded and the identity of individuals participating will remain confidential throughout the study.

The researcher will be available during and following the procedures in order to answer any inquiries concerning the procedures and to ensure that they are fully understood. Following completion of the study the researcher will be available for discussion or presentation of results and implications as well as to provide any further details regarding procedures and methodology.

CONSENT

It is the right of any subject to refuse to participate or to withdraw from the project at any time. Such a decision will neither jeopardize nor influence you in any way. Please indicate your willingness to participate in the project by providing your signature consenting to your participation.

I consent to participate in this project.

Signature

Date

PERSONAL DATA:

Gender M / F

Age

Highest level of education attained

Teaching experience: practicum
 years full time

Grade levels taught mainly

Have you had special training in child management?

If yes, please indicate with a check () below.

Courses Workshops

Inservice training

Other (please specify)

APPENDIX B

Knowledge Measure

KBPAC 16E

This measure consists of 16 items and will take approximately 8 to 10 minutes to complete.

Directions

1. You will be required to select responses to questions dealing with some typical child behaviors.
2. Read carefully each question and each of its four possible answers. In some questions and under certain circumstances more than one answer might be correct, however, you must select the single best answer or one which you feel is most generally true. Circle the letter beside that answer.

Example:

Probably the most important influence in a young child's life is his . . .

- a. Toys
- b. Television
- c. Parents
- d. Friends

Please do not consult others while deciding how to answer the questions. Complete all questions, even if you must guess.

KBPAC 16E

1. Which of the following is most important for parents in controlling their child's behavior?
 - a. The rules the parents make about behavior.
 - b. The parents' understanding of the child's feelings.
 - c. The behaviors to which the parents attend.
 - d. Being strict, but also warm and gentle.
2. A boy loves football. What is most likely to happen if, each time he is playing nicely with his sister, his father invites him to play football?
 - a. He will always be asking his father to play football.
 - b. He will play nicely with his sister more often.
 - c. He will be annoyed with his father for interfering with his activities.
 - d. He will be encouraged to teach his sister to play football.
3. If you want your child to develop proper study habits, you should:
 - a. Encourage him to do his homework.
 - b. Help him to see school as pleasant.
 - c. Reward him whenever he studies.
 - d. Give him good reasons why he will need school.

4. A child often cries over any small matter that bothers her. How should her parents react to best reduce her crying?
- Reward when she reacts without crying.
 - Use a mild punishment when she cries.
 - Try to find out what is really troubling the child and deal with that.
 - Provide her with something interesting so she will stop crying.
5. A father tells a child she cannot go to the store with him because she didn't clean her room like she promised. She reacts by shouting, crying and promising she will clean the room when she gets home. What should the father do?
- Ignore her and go to the store.
 - Take her to the store but make her clean her room when they return.
 - Calm her down and go help her clean her room together.
 - Talk to her and find out why she doesn't take responsibility.
6. If you want to make a behavior a long-lasting habit, you should:
- Reward it every time.
 - First reward it every time and then reward it occasionally.
 - Promise something the child wants very much.
 - Give several reasons why it is important and remind the child of the reasons often.

7. A baby often screams for several minutes and gets his parents' attention. Which of the following is probably the best way for his parents to reduce his screaming?
- a. If there is nothing physically wrong with the child, ignore his screaming even though the first few times he screams even louder.
 - b. Distract the child with something he finds interesting whenever he screams.
 - c. Ignore all noises and sounds the child makes.
 - d. None of the above. Babies usually have good reasons for screaming.
8. Probably the most important idea to keep in mind when first changing behavior is:
- a. To use both reward and punishment.
 - b. To reward every time the desired behavior occurs.
 - c. To be flexible about whether or not you reward.
 - d. To be sure the child understands why you want the behavior to change.
9. A child begins to whine and cry when his parent explains why he can't go outside. How should the parent react?
- a. Ask the child why going outside is so important to him.
 - b. Explain that it is a parent's right to make such decisions.
 - c. Explain again why he should not go outside.
 - d. Ignore the whining and crying.

10. If punishment is used for a behavior such as playing football in the house, which type is probably the best to use?

- a. Make the child do extra homework.
- b. Clearly express your disapproval.
- c. Remove the child to a boring situation each time.
- d. A reasonable spanking.

11. Which of the following is most effective in getting a child to do homework?

- a. "When you finish your homework, you can watch T.V."
- b. "You can watch this show on T.V., if you promise to do your homework when the show is over."
- c. "If you don't do your homework tonight, you can't watch T.V. at all tomorrow."
- d. Explain the importance of school work and the dangers of putting things off.

12. Parents who use lots of rewards for good behavior and few punishments will probably tend to have children who:

- a. Do not understand discipline.
- b. Will not cooperate unless they are 'paid'.
- c. Take advantage of their parents.
- d. Are well-behaved and cooperative.

13. Which reward is probably best to help a 12-year-old child improve his arithmetic skills?

- a. A dollar for each evening he studies.
- b. A dime for each problem he works correctly.
- c. Ten dollars for each A he receives on his report card in arithmetic.
- d. A bicycle for passing arithmetic for the rest of the year.

14. Johnny has just torn up a new magazine. Of the following choices, which is the best way for his mother to discipline him?

- a. Tell him he will be spanked by his father when he gets home.
- b. Punish him then and there.
- c. Explain to Johnny the wrongness of his actions.
- d. Angrily scold Johnny so that he will learn that such an act is bad and upsetting to his mother.

15. Jimmy sometimes says obscene words, but only in front of his mother. She has been shocked and makes her feelings clear to him. How should she react when he uses obscene words?

- a. Wash his mouth out with soap.
- b. Ignore him when he uses obscene words.
- c. Tell him how bad he is and how she doesn't like him when he uses those words.
- d. Explain to him the reason such words are not used.

16. Which of the following is probably most important in helping a child behave in desirable ways?

- a. To teach him the importance of self-discipline.
- b. To help him understand right and wrong.
- c. Providing consistent consequences for his behavior.
- d. Understanding his moods and feelings as a unique person.

APPENDIX C

Treatment Evaluation Inventory (Kazdin, 1980a)

Please complete the items listed below. The items should be completed by placing a checkmark on the line under the question that best indicates how you feel about the treatment. Please read the items very carefully because a checkmark accidentally placed on one space rather than another may not represent the meaning you intended.

Note that throughout the inventory the term treatment is used, and may be thought of as similar to such terms as intervention, method, procedure, or strategy.

1. How acceptable do you find this treatment to be for the child's problem behavior?

_____	_____	_____	_____
not at all		moderately	very
acceptable		acceptable	acceptable

2. How willing would you be to carry out this procedure yourself if you had to change the child's problems?

_____	_____	_____	_____
not at all		moderately	very
willing		willing	willing

3. How suitable is this procedure for children who might have other behavioral problems than those described for this child?

_____	_____	_____	_____
not at all		moderately	very
suitable		suitable	suitable

4. If children had to be assigned to treatment without their consent, how bad would it be to give them this treatment?

_____	_____	_____
very bad	moderately bad	not bad
		at all

5. How cruel or unfair do you find this treatment?

_____	_____	_____
very cruel	moderately cruel	not cruel at all

6. Would it be acceptable to apply this procedure to institutionalized children, the mentally retarded, or other individuals who are not given an opportunity to choose treatment for themselves?

_____	_____	_____
not at all acceptable to apply this procedure	moderately acceptable	very acceptable to apply this procedure

7. How consistent is this treatment with common sense or everyday notions about what treatment should be?

_____	_____	_____
very different or inconsistent	moderately consistent	very consistent with everyday notions

8. To what extent does this procedure treat the child humanely?

_____	_____	_____
does not treat humanely at all	treats them moderately humanely	treats them very humanely

9. To what extent do you think there might be risks in undergoing this kind of treatment?

_____	_____	_____
lots of risks are likely	some risks are likely	no risks are likely

10. How much do you like the procedures used in this treatment?

_____ do not like _____ moderately like _____ like them
them at all them very much

11. How effective is this treatment likely to be?

_____ not at all _____ moderately _____ very
effective effective effective

12. How likely is this treatment to make permanent improvements in the child?

_____ unlikely _____ moderately _____ very likely
likely

13. To what extent are undesirable side effects likely to result from this treatment?

_____ many _____ some _____ no _____
undesirable undesirable undesirable
side effects side effects side
likely likely effects
would occur

14. How much discomfort is the child likely to experience during the course of treatment?

_____ very much _____ moderate _____ no _____
discomfort discomfort discomfort at all

15. Overall, what is your general reaction to this form of treatment?

_____ very negative _____ ambivalent _____ very
positive

APPENDIX D

Semantic Differential (Kazdin, 1980a)

Please place your checkmarks on the line that best characterizes your reaction to the treatment. If the scale is difficult to rate, still put a checkmark that best reflects your general reaction to the treatment. There is no need to spend much time on any one of the items. Your first impressions and immediate feeling about the items are what I would like.

good	___	:	___	:	___	:	___	:	___	:	___	:	___	:	bad
pleasant	___	:	___	:	___	:	___	:	___	:	___	:	___	:	unpleasant
kind	___	:	___	:	___	:	___	:	___	:	___	:	___	:	cruel
valuable	___	:	___	:	___	:	___	:	___	:	___	:	___	:	worthless
fair	___	:	___	:	___	:	___	:	___	:	___	:	___	:	unfair
strong	___	:	___	:	___	:	___	:	___	:	___	:	___	:	weak
hard	___	:	___	:	___	:	___	:	___	:	___	:	___	:	soft
heavy	___	:	___	:	___	:	___	:	___	:	___	:	___	:	light
large	___	:	___	:	___	:	___	:	___	:	___	:	___	:	small
thick	___	:	___	:	___	:	___	:	___	:	___	:	___	:	thin
active	___	:	___	:	___	:	___	:	___	:	___	:	___	:	passive
sharp	___	:	___	:	___	:	___	:	___	:	___	:	___	:	dull
hot	___	:	___	:	___	:	___	:	___	:	___	:	___	:	cold
fast	___	:	___	:	___	:	___	:	___	:	___	:	___	:	slow
ferocious	___	:	___	:	___	:	___	:	___	:	___	:	___	:	peaceful

APPENDIX E

Case and Treatment Questionnaire

Now that you have had a chance to think about Mike's problems and about some possible strategies for dealing with Mike, I would like you to consider a few additional questions.

1. How serious would you rate the problems Mike is described as exhibiting?

5	4	3	2	1
extremely serious		moderately serious		not very serious at all

2. In your own experience, how typical is Mike of the kind of child being seen in the regular classroom?

3. Consider the four intervention strategies described for dealing with Mike's problems.

Positive Practice PP

Medication M

Reinforcement R

Time Out TO

Please rank these four interventions according to how likely you would be to use each in your classroom. Do so by placing the letter(s) corresponding to each intervention in the order which best represents your feelings.

least
likely to
use this
one

most likely
to use this
one

4. After reading through Mike's case description and the intervention strategies, you probably have had some reactions (e.g. "My God, who would ever do that to a kid", or, "Well of course, that's what I do all the time for kids like him!"). Teachers face this and similar kinds of problems in many classrooms. The reactions of those teachers to these kinds of situations are an important area of investigation.

In a few sentences, please describe how you might approach the situation if Mike were in your class.

APPENDIX F

Case Description

Case of Mike R.

Mike is a 9-year old boy who lives with his parents and younger brother, Tom, who is 6 years old. Although Mike is of normal intelligence and has no obvious academic difficulties, his social behavior has been a concern for some time.

His main problem is that he is very aggressive and does not follow adult instructions. He does many different things that are dangerous to himself or to others. Occasionally, he throws things at his teachers and his parents when they ask him to do things. He has broken many objects at school, such as books, chairs, toys, and a tape recorder, either by throwing them or knocking them off tables and shelves. He also has had problems at home. For example, he fights with his brother physically and sometimes throws sharp objects directly at him.

Both Mike's parents and his teacher are worried because they have not been able to control his aggressiveness. He has already injured his brother and his friends when he was angry, and most of the children in the school will not play with him because they are afraid he will hurt them. He has been in serious trouble in the community and at school. Mike's teacher is particularly concerned that things will get worse if something is not done right away.

Case of Mike S.

Mike is a 9-year old boy who lives with his parents and his younger brother, Tom, who is 6 years old. Although Mike is of normal intelligence and has no obvious academic difficulties, his social behavior has been of concern for some time. Mike has been in a special class managed by a teacher and her aide. In this setting, the teacher and aide conducted all activities during the day. The goals of this class were "to develop basic academic and social skills."

With the advent of program cuts and with a new emphasis in the school district on mainstreaming, Mike has been placed in a regular school setting. Unfortunately Mike is having some trouble in the class.

The main problem is that he is very aggressive and does not follow adult instructions. He does many different things that are dangerous to himself or to others. Occasionally, he throws things at his teachers and his parents when they ask him to do things. He has broken many objects at school, such as books, chairs, toys, and a tape recorder, either by throwing them or knocking them off tables and shelves. He also has had problems at home. For example, he fights with his brother physically and sometimes throws sharp objects directly at him.

Both Mike's parents and his teacher are worried because they have not been able to control his aggressiveness. He has already injured his brother and his friends when he was angry, and most of the children in the school will not play

with him because they are afraid he will hurt them. He has been in serious trouble in the community and at school. Mike's teacher is particularly concerned that things will get worse if something is not done right away.

APPENDIX G

Intervention Description

Medication Intervention for Mike

There is a drug that can be used to decrease aggressive behavior. The drug is called Mellaril. For aggressive children, the drug is used to increase a child's ability to tolerate frustration. While on the drug, children are able to handle events and interactions with others that ordinarily might be frustrating. Because the children do not become easily frustrated when on the drug, they are less likely to react aggressively to others. The drug seems to make some children less impulsive in general so that they are more able to think about the consequences of their actions before behaving inappropriately.

For Mike, the drug is administered in tablet form 3 times a day. The daily dosage of the drug is estimated on the basis of body weight and may be changed slightly until the optimal dose is found.

Time-Out Intervention for Mike

A procedure that can be used to change Mike's disruptive behavior is called TIME OUT FROM REINFORCEMENT. This procedure is designed to punish his aggressive and noncompliant behaviors so that they are eliminated. Time out means that there is a period of time in which a person is removed in some way from the situation. The procedure is called "time out from reinforcement" because during the time that Mike is taken away from the situation, he cannot receive reinforcement or rewards such as attention from others.

For Mike, the time out procedure consists of taking him out of the situation for 10 minutes whenever he misbehaves. For example, when Mike throws something or hits another pupil he is immediately taken out of the situation. He would be isolated in a part of the large classroom that is partitioned off for this purpose. The partition makes a little booth where Mike can be isolated. This time out booth consists of a chair in the booth. No books or toys are given to him so that he is cut off from all rewarding activities and from other people.

He has to sit in the booth for 10 minutes. He cannot see out nor can others see in because of the partition that closes him off from the rest of the classroom completely. After his time is up, the teacher lets him out. When Mike comes out, he can resume his activities. As soon as any aggression occurs, he is immediately placed into the time out booth again.

Reinforcement Intervention for Mike

A procedure that can be used to change Mike's aggressive behavior is called REINFORCEMENT OF INCOMPATIBLE BEHAVIOR. This is a procedure that is designed to develop compliant and cooperative behaviors to replace the aggressive behaviors. This is not a punishment procedure. It works, rather, by rewarding desirable behaviors. As the appropriate behaviors are developed, the inappropriate behaviors drop out and are eliminated.

Specifically, Mike's teacher will develop compliance to instructions and cooperative behavior by systematically rewarding appropriate responses. The teacher will explain the reward program to Mike. The program consists of letting Mike earn privileges if he listens to the teacher and behaves cooperatively. For example, if Mike follows directions in starting and completing his work on time, is cooperative in classroom routine and looks after his own and school property, he can earn Stars on a special chart. Mike can also earn Stars by playing nicely with his classmates and, when Mike get mad, he can earn Stars if he does not throw things, or does not act aggressively. The Stars can be traded for such privileges as extra recess, a special game, free time in class, or viewing an educational T.V. show for the class as a whole.

When Mike whines or argues the teacher will just ignore this. Mike will not earn Stars unless he behaves cooperatively.

Positive Practice Intervention for Mike

A procedure that can be used is called POSITIVE PRACTICE. This procedure is designed to suppress aggressive and noncompliant behaviors so that they are eliminated. Positive practice means that Mike repeatedly practices the correct behaviors to learn when he is supposed to do. When Mike misbehaves, he is immediately made to practice the appropriate behaviors that should have been done. He goes through the actual motions of appropriate behavior and, if necessary, is physically guided through the activities.

When Mike performs an aggressive act, the teacher comes over to him immediately. The teacher tells Mike he has to go through the appropriate behaviors. The teacher tells him exactly what to do and stays there while Mike goes through each of the appropriate behaviors. For example, when Mike throws something at a classmate, he is told to express his anger right then in a more appropriate manner. He must say he is mad and tell the other person why he is mad. Mike is told he must go through this sequence 3 times in a row to practice the correct behavior.

With positive practice, Mike practices expressing his anger and behaving appropriately rather than throwing things or hitting others. Repeatedly practicing the correct behaviors is an important part of the procedure.